

AN INVARIANT MANIFOLD THEORY FOR ODES AND ITS APPLICATIONS

DENNIS GUANG YANG

ABSTRACT. For a system of ODEs defined on an open, convex domain U containing a positively invariant set Γ , we prove that under appropriate hypotheses, Γ is the graph of a C^r function and thus a C^r manifold. Because the hypotheses can be easily verified by inspecting the vector field of the system, this invariant manifold theory can be used to study the existence of invariant manifolds in systems involving a wide range of parameters and the persistence of invariant manifolds whose normal hyperbolicity vanishes when a small parameter goes to zero. We apply this invariant manifold theory to study three examples and in each case obtain results that are not attainable by classical normally hyperbolic invariant manifold theory.

CONTENTS

1. Introduction	2
1.1. Main Results	3
1.2. The Problem of Weak Hyperbolicity	6
1.3. Organization	9
2. Proof of Theorem 1.1	9
2.1. The Existence and the Lipschitz Continuity of h	9
2.2. The C^1 Smoothness of h	11
3. Proof of Theorem 1.2	18
3.1. The Dynamics of Dh_0 along Trajectories in Γ	18
3.2. The C^2 Smoothness of h_0	19
3.3. The C^r Smoothness of h_0	22
4. Applications	25
4.1. Invariant Tori in a 2-Parameter Family of Systems	25
4.2. Invariant Torus in a System with Rapid Oscillations	30
4.3. Persistence of a Weakly Normally Hyperbolic Invariant Torus	35
Appendix A. The Wazewski Principle	45
References	46

Date: September, 2009.

2000 Mathematics Subject Classification. 37D10, 34C30.

Key words and phrases. invariant manifold, smoothness, weak hyperbolicity, invariant cone, the Wazewski principle.

1. INTRODUCTION

Consider the following ordinary differential equation

$$(1.1) \quad \dot{x} = F(x),$$

where $\dot{\cdot} = \frac{d}{dt}$, $x \in \mathbb{R}^n$, and F is C^r for some $r \geq 1$. Let $\Phi(t, x)$ be the flow generated by (1.1). Suppose M is a C^r invariant manifold of (1.1). Following the work of Fenichel [9], a simple version of normally hyperbolic invariant manifold theory can be roughly stated as follows: *if there is a continuous splitting of the tangent bundle of \mathbb{R}^n restricted to M : $T\mathbb{R}^n|_M = TM \oplus N^s \oplus N^u$ such that (1) $TM \oplus N^s$ and $TM \oplus N^u$ are invariant under $D_x\Phi(t, M)$ (i.e., the linearization of $\Phi(t, x)$ at M) and (2) $D_x\Phi(t, M)$ expands N^u and contracts N^s at rates at least r times of its expansion or contraction rate in TM (r -normal hyperbolicity), then M has a C^r stable manifold $W^s(M)$ tangent to N^s along M and a C^r unstable manifold $W^u(M)$ tangent to N^u along M , and the manifolds M , $W^s(M)$, and $W^u(M)$ all persist with the same C^r smoothness under any sufficiently small (in C^1 norm) C^r perturbation of the vector field F .* Here, we have omitted some technicalities such as the overflowing or inflowing invariance of M (if it has a boundary) and the local invariance of $W^s(M)$ and $W^u(M)$. We refer readers to the work of Fenichel [9] and an extensive exposition by Wiggins [19] for the precise description of the theory and many other properties of normally hyperbolic invariant manifolds. Equivalent results can also be found in the work of Hirsch, Pugh, and Shub [14]. A generalization to the case that M is an invariant set of (1.1) is given by Chow, Liu, and Yi in [6], where the authors proved the existence of center-stable, center-unstable, and center manifolds of M and their persistence under small perturbations. In [3], Bates, Lu, and Zeng developed a normally hyperbolic invariant manifold theory for C^1 semiflows in general Banach spaces. They also provided a thorough review on the extensive history of invariant manifold theory.

The persistence of normally hyperbolic invariant manifolds can be applied to establish the existence of invariant manifolds in systems in the form of

$$(1.2) \quad \dot{x} = F_0(x) + \epsilon F_1(x, \epsilon)$$

provided that for the corresponding “unperturbed” system $\dot{x} = F_0(x)$, the existence of an “unperturbed” normally hyperbolic C^r invariant manifold M_0 is already known. In particular, if M_0 is r -normally hyperbolic and the C^1 norm of the function $x \mapsto F_1(x, \epsilon)$ is $\mathcal{O}(\epsilon)$ for $\epsilon \rightarrow 0^+$, then a C^r invariant manifold M_ϵ , which is the perturbed counterpart of M_0 , exists for (1.2) with any $\epsilon \in (0, \epsilon_0]$ provided that ϵ_0 is sufficiently small. However, it can be a difficult task to verify the normal hyperbolicity of M_0 since the precise knowledge about M_0 and the linearization of the unperturbed flow at M_0 is not attainable in many applications. Furthermore, because normally hyperbolic invariant manifold theory does not provide any further information about the size of ϵ_0 other than being sufficiently small, we do not know to what extent the persistence result holds.

In this paper, we present a new theory on the existence of C^r invariant manifolds in systems of ordinary differential equations. Two special features of this new invariant manifold theory are: its hypotheses can be verified by inspecting the vector field instead of the flow of a specific system; and for systems depending on parameters, it provides a feasible way to compute the parameter ranges in which the desired invariant manifold results can be guaranteed. With these features, this new

theory can work with systems that require delicate analysis of the perturbations involved as well as systems that cannot be treated as perturbation problems, and in both cases, it establishes results that are not attainable by classical invariant manifold theory.

1.1. Main Results. We consider the following system

$$(1.3) \quad \begin{aligned} \dot{a} &= f(a, z), \\ \dot{z} &= g(a, z), \end{aligned}$$

where $\dot{\cdot} = \frac{d}{dt}$ and $(a, z) \in \mathbb{R}^n \times \mathbb{R}^m$. For notational convenience, we define $x := (a, z)$ and $X := \mathbb{R}^n \times \mathbb{R}^m$. For any function defined on a subset of X , we use x and (a, z) interchangeably to denote its argument, e.g., $f(x) = f(a, z)$, and $g(x) = g(a, z)$.

Suppose that f and g are at least C^1 on an open domain $U \subset \mathbb{R}^n \times \mathbb{R}^m$. Then (1.3) generates a flow $\Phi(t, x)$ on U even though for some $x \in U$, $\Phi(t, x)$ may not be defined for all $t \in \mathbb{R}$. For a subset of U , we define its positive invariance under the flow Φ as follows:

Definition (Positive Invariance). $\Gamma \subseteq U$ is positively invariant under Φ if

- (1) for every $x \in \Gamma$, $\Phi(t, x)$ is defined for all $t \geq 0$; and
- (2) $\Phi(t, \Gamma) \subseteq \Gamma$ for all $t \geq 0$.

First, we consider under what circumstances a positively invariant set Γ is a C^1 manifold. Let $\langle \cdot, \cdot \rangle$ denote the usual Euclidean inner product, and let $\|\cdot\|$ denote the usual Euclidean norm as well as the induced operator norm. For each $x = (a, z) \in X$ and $d > 0$, let $\mathbb{B}_d(x)$ be a closed box neighborhood of x :

$$\mathbb{B}_d(x) := \{(a', z') \in X : \|a' - a\| \leq d, \|z' - z\| \leq d\}.$$

In addition, define $\mathcal{L} : X \times X \rightarrow \mathbb{R}$ as follows:

$$(1.4) \quad \mathcal{L}(x_1, x_2) := \|a_2 - a_1\|^2 - \|z_2 - z_1\|^2,$$

where $x_1 = (a_1, z_1)$ and $x_2 = (a_2, z_2)$. Then for each $x \in X$, we define the cone with vertex x as follows:

$$(1.5) \quad \mathcal{C}(x) := \{x' \in X : \mathcal{L}(x', x) \geq 0\}.$$

Hypothesis 1. U is an open, convex subset of $\mathbb{R}^n \times \mathbb{R}^m$. In addition, there exists a $d > 0$ such that $\mathcal{C}(x) \cap U \subset \mathbb{B}_d(x)$ for all $x \in U$.

Hypothesis 2. f and g are C^1 on U . Furthermore, there exist a continuous, positive function $\alpha : U \rightarrow \mathbb{R}$, a continuous, nonnegative function $\ell : U \rightarrow \mathbb{R}$, and a constant $c_1 > 0$ such that

$$(1.6a) \quad \langle a', D_a f(x) a' \rangle \geq \alpha(x) \|a'\|^2 \text{ for any } x \in U \text{ and any } a' \in \mathbb{R}^n,$$

$$(1.6b) \quad \langle z', D_z g(x) z' \rangle \leq \ell(x) \|z'\|^2 \text{ for any } x \in U \text{ and any } z' \in \mathbb{R}^m,$$

$$(1.6c) \quad \alpha(x) \geq \ell(x) + \|D_z f(x)\| + \|D_a g(x)\| + c_1 \text{ for all } x \in U.$$

Define the following projections:

$$\begin{aligned} \Pi : (a, z) &\mapsto a, \\ \Pi_\perp : (a, z) &\mapsto z. \end{aligned}$$

Hypothesis 3. $\Gamma \subset U$ is positively invariant under the flow of (1.3) and satisfy $\Pi_\perp(\Gamma) = \Pi_\perp(U) \subseteq \mathbb{R}^m$.

Theorem 1.1. *Suppose Hypotheses 1, 2, and 3 hold. Let $K := \Pi_{\perp}(\Gamma) = \Pi_{\perp}(U)$. Then Γ contains all positively invariant subsets of U , and there exists a C^1 function $h : K \rightarrow \mathbb{R}^n$ such that $\Gamma = \{(h(z), z) : z \in K\}$. Moreover, $\|h(z_2) - h(z_1)\| < \|z_2 - z_1\|$ for any $z_1, z_2 \in K$ with $z_1 \neq z_2$, and $\|Dh(z)\| < 1$ for all $z \in K$.*

Next, we consider the further smoothness of a positively invariant set Γ if we know *a priori* that Γ is the graph of a C^1 function from an open subset of \mathbb{R}^m to \mathbb{R}^n .

Hypothesis 4. *K_0 is an open, convex subset of \mathbb{R}^m . $h_0 : K_0 \rightarrow \mathbb{R}^n$ is C^1 , and there exists an $\eta > 0$ such that $\|Dh_0(z)\| < \eta$ for all $z \in K_0$.*

Hypothesis 5. *There exists an open neighborhood U_0 of $\{(h_0(z), z) : z \in K_0\}$ such that f and g are C^r ($r \geq 2$) with their first to r -th derivatives all bounded on U_0 . Furthermore, there exist a continuous, positive function $\alpha : U_0 \rightarrow \mathbb{R}$, a continuous, nonnegative function $\ell : U_0 \rightarrow \mathbb{R}$, and a constant $c_r > 0$ such that*

$$(1.7a) \quad \langle a', D_a f(x) a' \rangle \geq \alpha(x) \|a'\|^2 \text{ for any } x \in U_0 \text{ and any } a' \in \mathbb{R}^n,$$

$$(1.7b) \quad \langle z', D_z g(x) z' \rangle \leq \ell(x) \|z'\|^2 \text{ for any } x \in U_0 \text{ and any } z' \in \mathbb{R}^m,$$

$$(1.7c) \quad \alpha(x) \geq r\ell(x) + (r+1)\eta\|D_a g(x)\| + c_r \text{ for all } x \in U_0.$$

Theorem 1.2. *Suppose Hypotheses 4 and 5 hold. If $\Gamma = \{(h_0(z), z) : z \in K_0\}$ is positively invariant under the flow of (1.3), then $h_0 : K_0 \rightarrow \mathbb{R}^n$ is C^r with its first to r -th derivatives all bounded on K_0 .*

It is evident that the combination of Theorem 1.1 and Theorem 1.2 establishes the existence of a C^r positively invariant manifold for (1.3). Specifically, take $U_0 = U$, $h_0 = h$, and $\eta = 1$. Then by replacing Hypothesis 2 with a stronger one, we obtain a C^r manifold theorem.

Hypothesis 2*. *f and g are C^r ($r \geq 2$) with their first to r -th derivatives all bounded on U . Furthermore, there exist a continuous, positive function $\alpha : U \rightarrow \mathbb{R}$, a continuous, nonnegative function $\ell : U \rightarrow \mathbb{R}$, and constants $c_1, c_r > 0$ such that*

$$(1.6d) \quad \begin{aligned} &\langle a', D_a f(x) a' \rangle \geq \alpha(x) \|a'\|^2 \text{ for any } x \in U \text{ and any } a' \in \mathbb{R}^n, \\ &\langle z', D_z g(x) z' \rangle \leq \ell(x) \|z'\|^2 \text{ for any } x \in U \text{ and any } z' \in \mathbb{R}^m, \\ &\alpha(x) \geq \ell(x) + \|D_z f(x)\| + \|D_a g(x)\| + c_1 \text{ for all } x \in U, \\ &\alpha(x) \geq r\ell(x) + (r+1)\|D_a g(x)\| + c_r \text{ for all } x \in U. \end{aligned}$$

Theorem 1.3. *Suppose Hypotheses 1, 2*, and 3 hold. Let $K := \Pi_{\perp}(\Gamma) = \Pi_{\perp}(U)$. Then Γ contains all positively invariant subsets of U , and there exists a C^r function $h : K \rightarrow \mathbb{R}^n$ such that $\Gamma = \{(h(z), z) : z \in K\}$. Moreover, $\|h(z_2) - h(z_1)\| < \|z_2 - z_1\|$ for any $z_1, z_2 \in K$ with $z_1 \neq z_2$, $\|Dh(z)\| < 1$ for all $z \in K$, and all higher derivatives of h are bounded on K .*

To apply Theorem 1.1 or Theorem 1.3 to establish invariant manifolds for a given system of ordinary differential equations, we need to carry out the following five steps. The first step is to reformulate the system into the form of (1.3). It is often necessary to change variables and append auxiliary parameters. In particular, appropriate rescaling of variables can lead to sharper results from the inequalities (1.6c) and (1.6d). The second step is to select a domain U that satisfies Hypothesis 1. Sometimes, it is necessary to choose a family of candidate domains that depends

on some parameters and then determine the parameters or their ranges in later analysis. The third step is to verify Hypothesis 3. Very often we define $\Gamma \subset U$ to be the set of points whose images under the flow of the reformulated system stay in U forever in forward time so that Γ is positively invariant by definition. In addition, this definition of Γ allows us to verify $\Pi_{\perp}(\Gamma) = \Pi_{\perp}(U)$ by simply inspecting the topological properties of the vector field at the boundary of U and then applying an elementary topological argument—the Ważewski principle (see Appendix A). The fourth step is to verify Hypothesis 2 or 2*. Especially, using the inequalities (1.6c) and (1.6d), we can estimate the set of admissible parameter values for the system or derive concrete, computable criteria for the smallness of perturbations. In the last step, we switch from the reformulated system back to the original system and identify the manifold that corresponds to Γ . Another advantage of the definition of Γ introduced in the third step is that it allows us to easily establish important properties, which may include uniqueness, full invariance (both forward time and backward time), independence of the rescaling of variables in the reformulation of the system (in the first step), periodicity with respect to some variables, etc., for the manifold in the original system. In Section 4, we will illustrate all these five steps and the relevant technical arguments in detail with three examples. Specifically, the first example is not a perturbation problem and involves a wide range of parameter values, and the other two examples are related to the problem of weak hyperbolicity, which will be discussed in the next subsection.

Traditionally, proofs of invariant manifold theorems are based on either one of two complementary methods: Hadamard’s graph transform method, and the Liapunov-Perron method. Both methods require the construction of a contraction in some Banach space such that its fixed point is a function whose graph is a Lipschitz invariant manifold. However, their approaches to constructing such a contraction are different. Specifically, the graph transform method obtains a contraction using the invariance of the manifold’s graph representation under a time- T map generated by the flow, whereas the Liapunov-Perron method achieves a contraction by setting up an integral equation using the variation of parameters formula. To prove the C^1 smoothness of the manifold, one constructs another contraction by formally differentiating the corresponding functional equations and then proves that its fixed point is in fact the desired derivative. If further smoothness is needed, one repeats this procedure inductively to show that the manifold is C^r smooth.

Alternatively, without referring to any contraction, one can establish invariant manifold results through a mixed use of invariant cones and topological arguments, an idea which, according to Jones [15], was introduced by Charles Conley. In [15], Jones used this method to construct a Lipschitz invariant manifold in a singularly perturbed slow-fast system. He also gave a very brief outline for showing the C^1 smoothness of the manifold. In an earlier reference [18], based on the same mixed use of invariant cones and topological arguments, McGehee supplied a proof of a C^r local stable manifold theorem for a fixed point of a hyperbolic linear map plus a nonlinear term that is small in C^r norm. In addition, Bates and Jones [2] extended this method to an infinite-dimensional setting to construct Lipschitz invariant manifolds for a semilinear partial differential equation.

In this paper, the proof of Theorem 1.1, which consists of two parts—the Lipschitz smoothness and the C^1 smoothness of Γ , is based on the mixed use of invariant cones and topological arguments. In fact, the Lipschitz smoothness part of the

proof closely follows the relevant part of the proof given in [15]. However, in the C^1 smoothness part of the proof, we introduce a new strategy for combining invariant cones and topological arguments. Specifically, we first construct a vector bundle over Γ such that it is invariant under the linearized dynamics along Γ and also satisfies some desirable topological and dynamical properties, and then we show that this vector bundle is in fact the tangent bundle of Γ . This approach is completely different from what is described in the outline for showing C^1 smoothness in [15] and what is given in the relevant part of the proof in [18]. The main technical issue is that the strategies used in [15, 18] rely on the assumption of “sufficient” C^r -smallness of certain terms (or perturbations) and thus are not applicable in our case, where the corresponding boundedness conditions are given explicitly by the inequality (1.6c).

In the proof of Theorem 1.2 for the C^r smoothness of Γ , we also introduce a new inductive scheme, which utilizes the hypothesis that the manifold is known to be the graph of the C^1 function h_0 . We show that in an appropriate coordinate system, along any solution trajectory of (1.3) on Γ the dynamics of the derivative Dh_0 are governed by a system in the form of (1.3). This allows us to apply Theorem 1.1 to show that Dh_0 is C^1 and thus h_0 is C^2 . Once again, by choosing an appropriate coordinate system, we can express the dynamics (along any solution trajectory of (1.3) on Γ) of the second derivative D^2h_0 by a system in the form of (1.3). Then we proceed inductively to establish the C^r smoothness of h_0 .

1.2. The Problem of Weak Hyperbolicity. Although the invariant manifold theory presented in this paper has a wide range of applications, the initial motivation for this work is to study the persistence of invariant manifolds whose normal hyperbolicity is “weak” in the sense that it depends on a small parameter and vanishes as the parameter goes to zero. Specifically, consider the system

$$(1.8) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w) + \epsilon^2 F_2(w, \theta, \epsilon), \\ \dot{\theta} &= G_0(w) + \epsilon G_1(w) + \epsilon^2 G_2(w, \theta, \epsilon), \end{aligned}$$

where $w \in \mathbb{R}^n$, $\theta \in \mathbb{T}^m$, $0 < \epsilon \ll 1$, and all functions on the right-hand side are C^r ($r \geq 1$) with respect to their arguments. Systems in the form of (1.8) often arise in situations involving averaging. In particular, in view of the absence of θ in the $\mathcal{O}(1)$ and $\mathcal{O}(\epsilon)$ terms, (1.8) can be regarded as the result of a first-order averaging procedure applied to a near-integrable system which is formulated in action-angle variables. In this case, the equation $\dot{w} = \epsilon F_1(w)$ is usually referred to as the averaged equation, and studying the dynamics of it is the first step to understand the dynamics of the full system (1.8). Suppose a C^r invariant manifold $M \subset \mathbb{R}^n$ is identified for the averaged equation $\dot{w} = \epsilon F_1(w)$. It follows that $M \times \mathbb{T}^m$ is a C^r invariant manifold in the truncated system

$$(1.9) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w), \\ \dot{\theta} &= G_0(w) + \epsilon G_1(w). \end{aligned}$$

Then the immediate question is the persistence of $M \times \mathbb{T}^m$ in the full system (1.8), which includes the $\mathcal{O}(\epsilon^2)$ terms. However, this is a nontrivial problem even under the assumption that $M \times \mathbb{T}^m$ is normally hyperbolic with respect to (1.9) for any fixed $\epsilon \in (0, \epsilon_0]$ with ϵ_0 being some positive constant. In particular, when we consider any fixed $\epsilon \in (0, \epsilon_0]$, it is unclear whether or not the $\mathcal{O}(\epsilon^2)$ terms are “sufficiently small” in C^1 norm so that the aforementioned persistence of normally

hyperbolic invariant manifolds applies. On the other hand, as we reduce the C^1 norm of the $\mathcal{O}(\epsilon^2)$ terms by reducing ϵ , the “strength” of the normal hyperbolicity of $M \times \mathbb{T}^m$ with respect to (1.9) is also weakened since the hyperbolicity is generated by the $\mathcal{O}(\epsilon)$ term $\epsilon F_1(w)$. In the limit $\epsilon \rightarrow 0^+$, the normal hyperbolicity of $M \times \mathbb{T}^m$ with respect to (1.9) fails as $\epsilon F_1(w)$ vanishes. This situation is referred to as *weak hyperbolicity* of the manifold $M \times \mathbb{T}^m$ [11, 19]. We also say that $M \times \mathbb{T}^m$ is only *weakly normally hyperbolic* with respect to (1.9). Note that by rescaling θ and time, we can obtain a variant of (1.8) and correspondingly a variant of (1.9), with respect to which the normal hyperbolicity of $M \times \mathbb{T}^m$ no longer depends on ϵ . However, the problem remains as other forms of singularities occur after rescaling variables. See [5] for the related discussions.

In [19], Wiggins adapted a continuation argument, which was originally proposed by Kopell [17] for a different class of systems, to show the persistence of $M \times \mathbb{T}^m$ in (1.8) for the case that $m \geq 1$ and M is an attracting fixed point of $\dot{w} = \epsilon F_1(w)$. Specifically, fix $\epsilon > 0$, and rewrite (1.8) as a one-parameter family of systems

$$(1.10) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w) + \alpha^2 F_2(w, \theta, \alpha), \\ \dot{\theta} &= G_0(w) + \epsilon G_1(w) + \alpha^2 G_2(w, \theta, \alpha), \end{aligned}$$

where $\alpha \in [0, \epsilon]$ is an auxiliary parameter. When $\alpha = 0$, (1.10) reduces to (1.9), under the flow of which $M \times \mathbb{T}^m$ is normally hyperbolic by the assumption that M is an attracting fixed point of $\dot{w} = \epsilon F_1(w)$. Then the persistence of normally hyperbolic invariant manifolds implies that there exists an $\alpha_1 \in (0, \epsilon]$ such that for any $\alpha \in [0, \alpha_1]$, (1.10) has an invariant torus \mathcal{M}_α which is a perturbation of $\mathcal{M}_0 := M \times \mathbb{T}^m$. Note that if \mathcal{M}_{α_1} is normally hyperbolic with respect to (1.10) with $\alpha = \alpha_1$, then α_1 can be increased to some $\alpha_2 \in (\alpha_1, \epsilon]$ such that \mathcal{M}_α exists for any $\alpha \in [0, \alpha_2]$. Now the question is whether or not this process can be repeated so that the family of invariant tori $\{\mathcal{M}_\alpha\}$ can be continued for all $\alpha \in [0, \epsilon]$. In [19] (pp. 168–170), Wiggins derived upper estimates of the generalized Lyapunov-type numbers¹ of \mathcal{M}_α . Then he argued that for sufficiently small $\epsilon > 0$, those upper estimates are bounded below the required critical values uniformly for all $\alpha \in [0, \epsilon]$ and thus \mathcal{M}_α exists and remains normally hyperbolic for any $\alpha \in [0, \epsilon]$.

This continuation argument was then used as a general strategy for establishing the persistence of weakly normally hyperbolic invariant manifolds in several different model systems (see, e.g., [12, 11]). However, Chicone and Liu subsequently pointed out in [5] that this argument contains a conceptual gap because it fails to address the issue that *the uniform boundedness of the generalized Lyapunov-type numbers does not guarantee uniform normal hyperbolicity for a continuous family of normally hyperbolic invariant manifolds*. Referring to the example discussed above, the technical issue can be described as follows. By extending the interval $[0, \alpha_1]$ to $[0, \alpha_2]$ and continuing in the same way, we can obtain an increasing sequence

¹We refer readers to the references [9, 19] for the technical definitions of generalized Lyapunov-type numbers. Roughly speaking, under the linearized dynamics along a trajectory on the invariant manifold M , the generalized Lyapunov-type numbers measure the expansion and contraction rates in the bundles N^u and N^s and compare these rates with the expansion or contraction rate in the tangent bundle TM . Thus, the invariant manifold M is normally hyperbolic if (1) M is C^r smooth for some $r \geq 1$, (2) it admits the splitting $T\mathbb{R}^n|_M = TM \oplus N^s \oplus N^u$ which satisfies the required invariance condition, and (3) the generalized Lyapunov-type numbers of M are bounded below certain critical values. Note that (1) and (2) are required so that the generalized Lyapunov-type numbers of M can be defined.

$\{\alpha_1, \alpha_2, \alpha_3, \dots\}$ such that \mathcal{M}_α exists for any $\alpha \in [0, \alpha_i]$ with $i = 1, 2, 3, \dots$. However, if $\{\alpha_1, \alpha_2, \alpha_3, \dots\}$ converges to a limit $\alpha^* < \epsilon$, we can only conclude that \mathcal{M}_α exists for any $\alpha \in [0, \alpha^*)$. In order to continue the family of tori $\{\mathcal{M}_\alpha\}$ for $\alpha \geq \alpha^*$, we must establish the existence of a normally hyperbolic invariant torus \mathcal{M}_{α^*} for (1.10) with $\alpha = \alpha^*$. Unfortunately, even if the generalized Lyapunov-type numbers of \mathcal{M}_α are bounded below their critical values uniformly for all $\alpha \in [0, \alpha^*)$, the existence of \mathcal{M}_{α^*} cannot be guaranteed because \mathcal{M}_α may still lose normal hyperbolicity in the limit $\alpha \rightarrow \alpha^{*-}$ by losing smoothness locally² or losing the splitting $T(\mathbb{R}^n \times \mathbb{T}^m)|_{\mathcal{M}_\alpha} = T\mathcal{M}_\alpha \oplus N^s \oplus N^u$ locally³. In particular, since the generalized Lyapunov-type numbers only reflect the global characteristics of the linearized dynamics along the invariant manifold, their uniform boundedness alone is not sufficient to rule out these local “defects”. Note that although the expressions of the upper estimates used in the continuation argument may remain defined and bounded below the required critical values for $\alpha \geq \alpha^*$, they no longer have any significance since the corresponding generalized Lyapunov-type numbers are not even defined before we actually establish the existence of \mathcal{M}_α for $\alpha \geq \alpha^*$.

Technically, the presence of the conceptual gap in the continuation argument does not imply that the result achieved by this argument is necessarily false. However, the assertion in [19] that an attracting invariant torus \mathcal{M}_α exists for (1.10) with any $\alpha \in [0, \epsilon]$ if ϵ is sufficiently small is in fact untrue, and so is the claim that for a sufficiently small $\epsilon > 0$, certain upper estimates of the generalized Lyapunov-type numbers of \mathcal{M}_α are bounded below the required critical values uniformly for all $\alpha \in [0, \epsilon]$. Indeed, it is possible that the generalized Lyapunov-type numbers converge to their critical values at some $\alpha < \epsilon$ no matter how small ϵ is. To see this, consider the following simple example, which is a particular case of (1.8):

$$(1.11) \quad \begin{aligned} \dot{w} &= -\epsilon w + \alpha^2 \sin \theta, \\ \dot{\theta} &= w, \end{aligned}$$

where $w \in \mathbb{R}$, $\theta \in \mathbb{T}^1$, $0 < \epsilon \ll 1$, and $\alpha \in [0, \epsilon]$. Clearly, for any fixed $\epsilon > 0$, $w = 0$ is an attracting fixed point of $\dot{w} = -\epsilon w$, and $\mathcal{M}_0 = \{(0, \theta) : \theta \in \mathbb{T}^1\}$ is a normally hyperbolic invariant 1-torus for (1.11) with $\alpha = 0$. However, no matter how small $\epsilon > 0$ is, (1.11) with $\alpha = \epsilon$ has two fixed points, which are a saddle point at $(0, 0)$ and a stable spiral at $(0, \pi)$, and thus possesses no invariant 1-torus. Furthermore, in this example the generalized Lyapunov-type numbers of \mathcal{M}_α do reach their critical values as α converges to $\frac{1}{2}\epsilon$ from below, and the family of C^1 1-tori $\{\mathcal{M}_\alpha\}$ only exists for $\alpha \in [0, \frac{1}{2}\epsilon)$.

The above counterexample also suggests that we should not expect the general persistence of weakly normally hyperbolic invariant manifolds in (1.8) and instead we should formulate additional hypotheses about the system in order to establish the desired invariant manifold results. For example, let M be a hyperbolic fixed

²In [5], Chicone and Liu presented a scenario that a family of normally hyperbolic limit cycles, with their generalized Lyapunov-type numbers uniformly bounded below the required critical values, cannot be continued further as the family converges to a nonsmooth homoclinic loop.

³In [13], Haro and de la Llave studied continuation of invariant tori for quasi-periodic perturbations of the standard map. They numerically observed and analyzed a situation that a one-parameter family of normally hyperbolic invariant 1-tori, with their generalized Lyapunov-type numbers uniformly bounded below the required critical values, cannot be continued further as the subbundles N^s and N^u converge locally when the continuation parameter approaches a critical value. We remark that similar examples can also be constructed for flows.

point or a hyperbolic periodic orbit of the averaged system $\dot{w} = \epsilon F_1(w)$. Suppose $\theta \in \mathbb{T}^1$ and $G_0(w) \neq 0$ for any $w \in M$. Then $M \times \mathbb{T}^1$ persists in (1.8) as a hyperbolic periodic orbit or a hyperbolic torus (see, e.g., [1, 4, 10]). In addition, Chow and Lu [7] studied a particular form of (1.8):

$$(1.12) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w) + \epsilon^2 F_2(w, \theta, \epsilon), \\ \dot{\theta} &= \Theta_0 + \epsilon G_1(w) + \epsilon^2 G_2(w, \theta, \epsilon), \end{aligned}$$

where $\theta \in \mathbb{T}^m$ and the frequency vector Θ_0 is constant. For the case that M is a fixed point of $\dot{w} = \epsilon F_1(w)$ with center-stable, center-unstable, and center manifolds $W^{cs}(M)$, $W^{cu}(M)$, and $W^c(M)$, respectively, they proved the persistence of $W^{cs}(M) \times \mathbb{T}^m$, $W^{cu}(M) \times \mathbb{T}^m$, and $W^c(M) \times \mathbb{T}^m$ in (1.12) for sufficiently small $\epsilon > 0$. In [5], Chicone and Liu studied a 3-dimensional system (see (4.10)) which, by rescaling an angular variable and time, can be put into the following form:

$$(1.13) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w) + \epsilon^{\frac{3}{2}} F_2(w, \theta, \epsilon^{\frac{1}{2}}), \\ \dot{\theta} &= 1, \end{aligned}$$

where $w \in \mathbb{R} \times \mathbb{T}^1$, $\theta \in \mathbb{T}^1$, and ϵ is the square of the original small parameter used in [5]. Chicone and Liu proved that if M is an attracting or repelling limit cycle of $\dot{w} = \epsilon F_1(w)$ then $M \times \mathbb{T}^1$ persists in (1.13) for sufficiently small $\epsilon > 0$. Note that the perturbation is $\mathcal{O}(\epsilon^{\frac{3}{2}})$ in this case.

In Subsection 4.3, we will study the following system

$$(1.14) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w) + \epsilon^{1+\mu} F_2(w, \theta, \epsilon), \\ \dot{\theta} &= \Theta_0 + \epsilon^\nu G_1(w, \epsilon) + \epsilon^{1+\gamma} G_2(w, \theta, \epsilon), \end{aligned}$$

where $w \in \mathbb{R}^n$, $\theta \in \mathbb{T}^m$, $\mu > 0$, $\gamma > 0$, $1 \geq \nu \geq 0$, and $\mu + \nu > 1$. Note that both (1.12) and (1.13) are special cases of (1.14). Specifically, we have $\mu = 1$, $\gamma = 1$, and $\nu = 1$ for (1.12) and $\mu = \frac{1}{2}$, $\gamma = 1$, and $\nu = 1$ for (1.13) (considering $\dot{\theta} = 1 + \epsilon\theta + \epsilon^2\theta$). We will apply Theorem 1.1 and Theorem 1.3 to show that if M is a hyperbolic periodic orbit of $\dot{w} = \epsilon F_1(w)$ then $M \times \mathbb{T}^m$ persists in (1.14) for sufficiently small $\epsilon > 0$ (see Theorem 4.4). The crucial difference between (1.14) (or its special cases (1.12) and (1.13)) and (1.11) with $\alpha = \epsilon$ is that for the latter case, $\mu = 1$ and $\nu = 0$ and thus the inequality $\mu + \nu > 1$ is not satisfied. Finally, we mention that in Subsection 4.2, we will study (1.13) in its original form in [5] and solve an open problem posed by Chicone and Liu. In particular, by applying Theorem 1.1 and Theorem 1.3, we will formulate a sufficient condition for the existence of a C^r invariant torus without assuming the existence of any unperturbed invariant torus.

1.3. Organization. The balance of this paper is organized as follows. The complete proofs of Theorem 1.1 and Theorem 1.2 are given in Section 2 and Section 3, respectively. Then we illustrate the applications of Theorem 1.1 and Theorem 1.3 with three examples in Section 4. At the end of this paper, a concise statement of the Ważewski principle is included in Appendix A.

2. PROOF OF THEOREM 1.1

2.1. The Existence and the Lipschitz Continuity of h . In this subsection, we prove that there is a Lipschitz function $h : K \rightarrow \mathbb{R}^n$ such that the graph of h is Γ . We achieve this goal in three steps. In the first step (Lemma 2.1), we establish the

“invariance” of the “moving cones” $\mathcal{C}(\Phi(t, x))$, which move in translation as their vertices move under the flow Φ in forward time. In the second step (Lemma 2.2), we show that trajectories in these cones drift away from the moving vertices at least exponentially fast in forward time. In the third step (Lemma 2.3), we establish the existence and the Lipschitz continuity of h using Lemmas 2.1 and 2.2. Throughout this subsection, we will use the following notations for any $x_1, x_2 \in U$:

$$\begin{aligned}(\tilde{a}_1(t), \tilde{z}_1(t)) &= \tilde{x}_1(t) := \Phi(t, x_1), \\(\tilde{a}_2(t), \tilde{z}_2(t)) &= \tilde{x}_2(t) := \Phi(t, x_2), \\(\delta a(t), \delta z(t)) &= \delta x(t) := \tilde{x}_2(t) - \tilde{x}_1(t).\end{aligned}$$

In addition, taking account of the convexity of U , we define the functions $L_{f,a}$, $L_{f,z}$, $L_{g,a}$, and $L_{g,z}$, all mapping $U \times U$ into \mathbb{R} , as follows:

$$\begin{aligned}(2.1) \quad L_{f,a}(x_1, x_2) &:= \int_0^1 \alpha(x_1 + s(x_2 - x_1)) ds, \\L_{f,z}(x_1, x_2) &:= \int_0^1 \|D_z f(x_1 + s(x_2 - x_1))\| ds, \\L_{g,a}(x_1, x_2) &:= \int_0^1 \|D_a g(x_1 + s(x_2 - x_1))\| ds, \\L_{g,z}(x_1, x_2) &:= \int_0^1 \ell(x_1 + s(x_2 - x_1)) ds.\end{aligned}$$

Then it follows from (1.6a) and (1.6b) of Hypothesis 2 that

$$\begin{aligned}(2.2) \quad \langle \delta a, \dot{\delta a} \rangle &\geq L_{f,a}(\tilde{x}_1, \tilde{x}_2) \|\delta a\|^2 - L_{f,z}(\tilde{x}_1, \tilde{x}_2) \|\delta a\| \|\delta z\|, \\ \langle \delta z, \dot{\delta z} \rangle &\leq L_{g,a}(\tilde{x}_1, \tilde{x}_2) \|\delta a\| \|\delta z\| + L_{g,z}(\tilde{x}_1, \tilde{x}_2) \|\delta z\|^2,\end{aligned}$$

whenever $\tilde{x}_1(t) \in U$ and $\tilde{x}_2(t) \in U$.

Lemma 2.1. *For any $x_1, x_2 \in U$ and $t_0 > 0$, if $x_2 \in \mathcal{C}(x_1)$, $\Phi([0, t_0], x_1) \subset U$, and $\Phi([0, t_0], x_2) \subset U$, then $\tilde{x}_2(t) \in \mathcal{C}(\tilde{x}_1(t))$ for all $t \in [0, t_0]$.*

Proof. The lemma is trivially true for $x_1 = x_2 \in U$. Thus, we only consider $x_1, x_2 \in U$ with $x_1 \neq x_2$. Then $\tilde{x}_1(t) \neq \tilde{x}_2(t)$ for all possible t due to the uniqueness property of the solutions of (1.3) inside U . Suppose that $\Phi([0, t_0], x_1) \subset U$ and $\Phi([0, t_0], x_2) \subset U$ for some $t_0 > 0$.

Assume that $\tilde{x}_2(\tau) \in \mathcal{C}(\tilde{x}_1(\tau))$ for some $\tau \in [0, t_0]$. Then $\|\delta z(\tau)\| \leq \|\delta a(\tau)\|$ by (1.4) and (1.5). Incorporate this relation into (2.2). It follows that

$$(2.3a) \quad \langle \delta a(\tau), \dot{\delta a}(\tau) \rangle \geq (L_{f,a}(\tilde{x}_1(\tau), \tilde{x}_2(\tau)) - L_{f,z}(\tilde{x}_1(\tau), \tilde{x}_2(\tau))) \|\delta a(\tau)\|^2,$$

$$(2.3b) \quad \langle \delta z(\tau), \dot{\delta z}(\tau) \rangle \leq (L_{g,a}(\tilde{x}_1(\tau), \tilde{x}_2(\tau)) + L_{g,z}(\tilde{x}_1(\tau), \tilde{x}_2(\tau))) \|\delta a(\tau)\|^2.$$

Combining the above two inequalities and using (2.1) and (1.6c), we obtain

$$\langle \delta a(\tau), \dot{\delta a}(\tau) \rangle - \langle \delta z(\tau), \dot{\delta z}(\tau) \rangle \geq c_1 \|\delta a(\tau)\|^2.$$

Thus, for any $\tau \in [0, t_0]$ such that $\tilde{x}_2(\tau) \in \mathcal{C}(\tilde{x}_1(\tau))$ (i.e., $\mathcal{L}(\tilde{x}_2(\tau), \tilde{x}_1(\tau)) \geq 0$), we have that

$$(2.4) \quad \left[\frac{d}{dt} \mathcal{L}(\tilde{x}_2(t), \tilde{x}_1(t)) \right]_{t=\tau} \geq 2c_1 \|\delta a(\tau)\|^2 > 0,$$

where the second (strict) inequality is assured by $\|\delta a(\tau)\| > 0$, a fact due to $\tilde{x}_1(\tau) \neq \tilde{x}_2(\tau)$ and $\|\delta a(\tau)\| \geq \|\delta z(\tau)\|$. Now consider $x_2 \in \mathcal{C}(x_1)$. Then

$$(2.5) \quad \mathcal{L}(\tilde{x}_2(0), \tilde{x}_1(0)) \geq 0,$$

and with $\tau = 0$, (2.4) becomes

$$(2.6) \quad \left[\frac{d}{dt} \mathcal{L}(\tilde{x}_2(t), \tilde{x}_1(t)) \right]_{t=0} > 0.$$

Altogether, (2.4)–(2.6) imply that $\mathcal{L}(\tilde{x}_2(t), \tilde{x}_1(t)) \geq 0$ for all $t \in [0, t_0]$ if $x_2 \in \mathcal{C}(x_1)$. \square

We now demonstrate that inside the moving cone $\mathcal{C}(\Phi(t, x))$, trajectories drift away from the moving vertex $\Phi(t, x)$ at least exponentially fast in forward time.

Lemma 2.2. *For any $x_1, x_2 \in U$ and $t_0 > 0$, if $x_2 \in \mathcal{C}(x_1)$, $\Phi([0, t_0], x_1) \subset U$, and $\Phi([0, t_0], x_2) \subset U$, then $\|\tilde{a}_2(t) - \tilde{a}_1(t)\| \geq \|a_2 - a_1\| e^{c_1 t}$ for all $t \in [0, t_0]$.*

Proof. Take $x_1, x_2 \in U$ and $t_0 > 0$ such that $\Phi([0, t_0], x_1) \subset U$ and $\Phi([0, t_0], x_2) \subset U$. Assume that $\tilde{x}_2(\tau) \in \mathcal{C}(\tilde{x}_1(\tau))$ for some $\tau \in [0, t_0]$. Recall (2.3a), and note that $L_{f,a}(x, x') - L_{f,z}(x, x') \geq c_1$ for any $(x, x') \in U \times U$ by (1.6c). Then we have

$$(2.7) \quad \left[\frac{d}{dt} \|\tilde{a}_2(t) - \tilde{a}_1(t)\|^2 \right]_{t=\tau} \geq 2c_1 \|\tilde{a}_2(\tau) - \tilde{a}_1(\tau)\|^2.$$

Now consider $x_2 \in \mathcal{C}(x_1)$. By Lemma 2.1, $\tilde{x}_2(\tau) \in \mathcal{C}(\tilde{x}_1(\tau))$ for all $\tau \in [0, t_0]$. Thus (2.7) holds for all $\tau \in [0, t_0]$. It follows that $\|\tilde{a}_2(t) - \tilde{a}_1(t)\| \geq \|a_2 - a_1\| e^{c_1 t}$ for all $t \in [0, t_0]$. \square

Having established Lemmas 2.1 and 2.2, we now show that Γ is the largest positively invariant subset of U and it is the graph of a Lipschitz function $h : K \rightarrow \mathbb{R}^n$.

Lemma 2.3. *Γ contains all positively invariant subsets of U , and there exists a Lipschitz function $h : K \rightarrow \mathbb{R}^n$ such that $\Gamma = \{(h(z), z) : z \in K\}$. Moreover, $\|h(z_2) - h(z_1)\| < \|z_2 - z_1\|$ for any $z_1, z_2 \in K$ with $z_1 \neq z_2$.*

Proof. Let Γ' be the union of all positively invariant subsets of U . Consider $x_1, x_2 \in \Gamma'$ with $x_1 \neq x_2$. First, we prove $x_2 \notin \mathcal{C}(x_1)$ by contradiction. Suppose $x_2 \in \mathcal{C}(x_1)$. Since Γ' is positively invariant, $\Phi([0, t_0], x_1) \subset \Gamma' \subseteq U$ and $\Phi([0, t_0], x_2) \subset \Gamma' \subseteq U$ for all $t_0 > 0$. Applying Lemma 2.2, we obtain that for all $t \geq 0$,

$$\|\tilde{a}_2(t) - \tilde{a}_1(t)\| \geq \|a_2 - a_1\| e^{c_1 t}.$$

In addition, $\|a_2 - a_1\| > 0$ since $x_1 \neq x_2$ and $\|a_2 - a_1\| \geq \|z_2 - z_1\|$ for $x_2 \in \mathcal{C}(x_1)$. Thus, $\|a_2 - a_1\| e^{c_1 t}$ increases exponentially (as opposed to being constant 0) for $t \geq 0$. On the other hand, Lemma 2.1 implies that $\tilde{x}_2(t) \in \mathcal{C}(\tilde{x}_1(t))$ for all $t \geq 0$. Then it follows from Hypothesis 1 that $\|\tilde{a}_2(t) - \tilde{a}_1(t)\| < d$ for all $t \geq 0$, which constitutes a contradiction.

It follows that for each $x \in \Gamma'$, $\mathcal{C}(x) \cap \Gamma' = \{x\}$. Since $K = \Pi_\perp(\Gamma) = \Pi_\perp(U)$ by Hypothesis 3 and $\Pi_\perp(\Gamma) \subseteq \Pi_\perp(\Gamma') \subseteq \Pi_\perp(U)$, we have that $\Pi_\perp(\Gamma') = K$. Then for each $z \in K$ there is a unique $x_z = (a_z, z) \in \Gamma'$ such that $\Pi_\perp(x_z) = z$. Thus, we can define a function $h : K \rightarrow \mathbb{R}^n$ by letting $h(z) := a_z$ for each $z \in K$. Obviously, $\Gamma' = \{(h(z), z) : z \in K\}$. In addition, we have $\Gamma = \Gamma'$ since $\Pi_\perp(\Gamma) = \Pi_\perp(\Gamma') = K$.

Next, we consider $x_1 = (h(z_1), z_1)$ and $x_2 = (h(z_2), z_2)$ for any $z_1, z_2 \in K$ with $z_1 \neq z_2$. Since $x_1, x_2 \in \Gamma$ and $x_1 \neq x_2$, we have $x_2 \notin \mathcal{C}(x_1)$. Then it follows from (1.4) and (1.5) that $\|h(z_2) - h(z_1)\| < \|z_2 - z_1\|$. \square

2.2. The C^1 Smoothness of h . We need to work with the variational equation of (1.3) along trajectories in Γ . Rewrite (1.3) in a compact form: $\dot{x} = F(x)$, and let $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{z} \in \mathbb{R}^m$, and $\mathbf{x} = (\mathbf{a}, \mathbf{z}) \in X$ be the variations of a , z , and x , respectively. Then the variational equation of (1.3) along a trajectory $\Phi(t, x)$ for any $x \in \Gamma$ is given by

$$(2.8) \quad \dot{\mathbf{x}} = DF(\Phi(t, x)) \mathbf{x},$$

where $t \geq 0$ and $x \in \Gamma$ now serves as a parameter.

For the linear system (2.8) with any parameter $x \in \Gamma$, the solution that originates at \mathbf{x} at $t = 0$ can be represented as $Q(t, x) \mathbf{x}$, where $Q(t, x)$ is a linear transformation of X to itself for each fixed $t \geq 0$ with $Q(0, x) = I$ being the identity transformation of X . Note that for f and g that are C^1 on U , $DF(\Phi(t, x))$ depends on (t, x) continuously on $[0, \infty) \times \Gamma$. Then for each fixed $x \in \Gamma$, $Q(t, x)$ is defined for all $t \geq 0$, and furthermore, the map $(t, x) \mapsto Q(t, x)$ is continuous on $[0, \infty) \times \Gamma$.

It is important to note that $Q(t, x) \mathbf{x}$ is *not* a flow on X since the system (2.8) is *nonautonomous*. However, for $(t, x) \in [0, \infty) \times \Gamma$, the family of linear transformations $Q(t, x)$ forms a *cocycle* over the flow $\Phi(t, x)$, i.e., for any $\tau, t \in [0, \infty)$ and $x \in \Gamma$,

$$\begin{aligned} Q(0, x) &= I, \\ (2.9) \quad Q(\tau + t, x) &= Q(\tau, \Phi(t, x)) \circ Q(t, x). \end{aligned}$$

Note that our construction of cones in X still applies when using the variational variable $\mathbf{x} = (\mathbf{a}, \mathbf{z})$, i.e.,

$$(2.10) \quad \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{a}_2 - \mathbf{a}_1\|^2 - \|\mathbf{z}_2 - \mathbf{z}_1\|^2,$$

$$(2.11) \quad \mathcal{C}(\mathbf{x}) = \{\mathbf{x}' \in X : \mathcal{L}(\mathbf{x}', \mathbf{x}) \geq 0\}.$$

Let $\mathbf{0}$ be the zero vector in X . For each $x \in \Gamma$, define the set $T(x)$ as follows:

$$\begin{aligned} (2.12) \quad T(x) &:= \{\mathbf{x} \in X : Q(t, x) \mathbf{x} \notin \text{int}(\mathcal{C}(\mathbf{0})) \text{ for all } t \geq 0\} \\ &= \{\mathbf{x} \in X : \mathcal{L}(Q(t, x) \mathbf{x}, \mathbf{0}) \leq 0 \text{ for all } t \geq 0\}. \end{aligned}$$

Obviously, $\mathbf{0} \in T(x)$, and the image of $T(x)$ under the linear transformation $Q(t, x)$ does not intersect $\text{int}(\mathcal{C}(\mathbf{0}))$ for any $t \geq 0$. In addition, it follows from (2.9) that for any $\mathbf{x} \in T(x)$,

$$(2.13) \quad Q(t, x) \mathbf{x} \in T(\Phi(t, x)) \text{ for all } t \geq 0.$$

The outline of the proof of the C^1 smoothness of h is the following. In 2.2.1, we will first show that for each $x \in \Gamma$, $T(x)$ is in fact a linear subspace of X and it can be represented as the graph of a linear operator $H(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$. Then we will demonstrate in 2.2.2 that the map $x \mapsto H(x)$ is continuous for all $x \in \Gamma$. Finally, in 2.2.3, we will show that $H(h(z), z)$ is indeed the derivative of h at z for any $z \in K$.

2.2.1. $T(x)$ is a linear subspace of X . Consider (2.8) in terms of (\mathbf{a}, \mathbf{z}) , i.e.,

$$\begin{aligned} (2.14) \quad \dot{\mathbf{a}} &= D_a f(\Phi(t, x)) \mathbf{a} + D_z f(\Phi(t, x)) \mathbf{z}, \\ \dot{\mathbf{z}} &= D_a g(\Phi(t, x)) \mathbf{a} + D_z g(\Phi(t, x)) \mathbf{z}. \end{aligned}$$

Applying (1.6a) and (1.6b), we obtain the following inequalities:

$$(2.15a) \quad \langle \mathbf{a}, \dot{\mathbf{a}} \rangle \geq \alpha(\Phi(t, x)) \|\mathbf{a}\|^2 - \|D_z f(\Phi(t, x))\| \|\mathbf{a}\| \|\mathbf{z}\|,$$

$$(2.15b) \quad \langle \mathbf{z}, \dot{\mathbf{z}} \rangle \leq \|D_a g(\Phi(t, x))\| \|\mathbf{a}\| \|\mathbf{z}\| + \ell(\Phi(t, x)) \|\mathbf{z}\|^2,$$

which are similar to (2.2). In addition, for any $x \in \Gamma$, both inequalities of (2.15) hold for all $t \geq 0$ since $\Phi(t, x) \in \Gamma \subset U$ for all $t \geq 0$.

The next three lemmas are simple consequences of (2.15).

Lemma 2.4. For any $x \in \Gamma$ and any $\mathbf{x} \in \mathcal{C}(\mathbf{0})$, $Q(t, x) \mathbf{x} \in \mathcal{C}(\mathbf{0})$ for all $t \geq 0$.

Lemma 2.5. *For any $x \in \Gamma$ and any $\mathbf{x} \in \mathcal{C}(\mathbf{0})$,*

$$\|\Pi(Q(t, x) \mathbf{x})\| \geq \|\Pi(\mathbf{x})\| e^{\int_0^t (\alpha(\Phi(s, x)) - \|D_z f(\Phi(s, x))\|) ds} \text{ for all } t \geq 0.$$

We omit the proofs of Lemma 2.4 and Lemma 2.5 since they are essentially the same as the proofs of Lemma 2.1 and Lemma 2.2, respectively. A difference here is that we only consider $x \in \Gamma$. Thus the statements of Lemma 2.4 and Lemma 2.5 can be shown true for all $t \geq 0$ (as opposed to only for $t \in [0, t_0]$).

By (1.6c), we have that for any $(t, x) \in [0, \infty) \times \Gamma$,

$$(2.16) \quad \alpha(\Phi(t, x)) - \|D_z f(\Phi(t, x))\| \geq c_1 + \|D_a g(\Phi(t, x))\| + \ell(\Phi(t, x)).$$

Thus Lemma 2.5 implies that for any $x \in \Gamma$ and $\mathbf{x} \in \mathcal{C}(\mathbf{0})$ with $\mathbf{x} \neq \mathbf{0}$, $\|\Pi(Q(t, x) \mathbf{x})\|$ and $\|Q(t, x) \mathbf{x}\|$ both grow at least exponentially fast as t increases.

Next, for $\mathbf{x} \in T(x)$, we have the following growth estimate of $\|Q(t, x) \mathbf{x}\|$.

Lemma 2.6. *For any $x \in \Gamma$ and any $\mathbf{x} \in T(x)$,*

$$\|Q(t, x) \mathbf{x}\| \leq 2\|\mathbf{x}\| e^{\int_0^t (\|D_a g(\Phi(s, x))\| + \ell(\Phi(s, x))) ds} \text{ for all } t \geq 0.$$

Proof. Take an arbitrary $x \in \Gamma$ and then an arbitrary $\mathbf{x} \in T(x)$. Recall (2.10) and (2.11) for $\mathcal{C}(\mathbf{0})$. By the definition of $T(x)$ (see (2.12)), we have that

$$(2.17) \quad \|\Pi(Q(t, x) \mathbf{x})\| \leq \|\Pi_\perp(Q(t, x) \mathbf{x})\| \text{ and}$$

$$(2.18) \quad \|Q(t, x) \mathbf{x}\| \leq 2\|\Pi_\perp(Q(t, x) \mathbf{x})\|$$

for all $t \geq 0$. Then it follows from (2.15b) and (2.17) that for all $t \geq 0$,

$$\frac{d}{dt} \|\Pi_\perp(Q(t, x) \mathbf{x})\|^2 \leq 2 \left(\|D_a g(\Phi(t, x))\| + \ell(\Phi(t, x)) \right) \|\Pi_\perp(Q(t, x) \mathbf{x})\|^2.$$

Thus, for all $t \geq 0$,

$$\|\Pi_\perp(Q(t, x) \mathbf{x})\| \leq \|\Pi_\perp(\mathbf{x})\| e^{\int_0^t (\|D_a g(\Phi(s, x))\| + \ell(\Phi(s, x))) ds}.$$

Then using (2.18), we obtain that for all $t \geq 0$,

$$\begin{aligned} \|Q(t, x) \mathbf{x}\| &\leq 2\|\Pi_\perp(\mathbf{x})\| e^{\int_0^t (\|D_a g(\Phi(s, x))\| + \ell(\Phi(s, x))) ds} \\ &\leq 2\|\mathbf{x}\| e^{\int_0^t (\|D_a g(\Phi(s, x))\| + \ell(\Phi(s, x))) ds}. \end{aligned} \quad \square$$

We are now ready to demonstrate that $T(x)$ is a linear subspace of X .

Lemma 2.7. *For each $x \in \Gamma$, $T(x)$ is a linear subspace of X .*

Proof. Take an arbitrary $x \in \Gamma$. Consider any $\mathbf{x}_1, \mathbf{x}_2 \in T(x)$ and any $\lambda_1, \lambda_2 \in \mathbb{R}$. We prove that $\mathbf{x}_3 = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \in T(x)$ by contradiction.

Suppose $\mathbf{x}_3 \notin T(x)$. Then there exists a $t_0 \geq 0$ such that $Q(t_0, x) \mathbf{x}_3 \in \text{int}(\mathcal{C}(\mathbf{0}))$. Applying Lemma 2.5 to $x' := \Phi(t_0, x) \in \Gamma$ and $\mathbf{x}' := Q(t_0, x) \mathbf{x}_3$, we obtain that for all $\tau \geq 0$,

$$\|\Pi(Q(\tau, x') \mathbf{x}')\| \geq \|\Pi(\mathbf{x}')\| e^{\int_0^\tau (\alpha(\Phi(s, x')) - \|D_z f(\Phi(s, x'))\|) ds},$$

which, by (2.9), can be rewritten in terms of x and \mathbf{x}_3 as follows:

$$\|\Pi(Q(\tau + t_0, x) \mathbf{x}_3)\| \geq \|\Pi(Q(t_0, x) \mathbf{x}_3)\| e^{\int_0^\tau (\alpha(\Phi(s + t_0, x)) - \|D_z f(\Phi(s + t_0, x))\|) ds}.$$

Let $t = \tau + t_0$. Then the above estimate becomes

$$(2.19) \quad \|\Pi(Q(t, x) \mathbf{x}_3)\| \geq \|\Pi(Q(t_0, x) \mathbf{x}_3)\| e^{\int_{t_0}^t (\alpha(\Phi(s, x)) - \|D_z f(\Phi(s, x))\|) ds}$$

for all $t \geq t_0$. Note that $\|\Pi(Q(t_0, x) \mathbf{x}_3)\| > 0$ since $Q(t_0, x) \mathbf{x}_3 \in \text{int}(\mathcal{C}(\mathbf{0}))$. Thus, the right-hand side of (2.19) increases at least exponentially (as opposed to being constant 0) for $t \geq t_0$.

On the other hand, it follows from Lemma 2.6 that for all $t \geq t_0$,

$$\begin{aligned} \|Q(t, x) \mathbf{x}_3\| &\leq |\lambda_1| \|Q(t, x) \mathbf{x}_1\| + |\lambda_2| \|Q(t, x) \mathbf{x}_2\| \\ &\leq 2(|\lambda_1| \|\mathbf{x}_1\| + |\lambda_2| \|\mathbf{x}_2\|) e^{\int_0^t (\|D_a g(\Phi(s, x))\| + \ell(\Phi(s, x))) ds} \\ &\leq \mu e^{\int_{t_0}^t (\|D_a g(\Phi(s, x))\| + \ell(\Phi(s, x))) ds}, \end{aligned}$$

where μ in the last inequality is a sufficiently large constant. By (2.16), this estimate contradicts (2.19). \square

Let $\mathcal{I}(\mathbf{z})$ denote the cross-section of X at \mathbf{z} along the \mathbf{a} -direction, i.e.,

$$\mathcal{I}(\mathbf{z}) := \{(\mathbf{a}, \mathbf{z}) : \mathbf{a} \in \mathbb{R}^n\}.$$

A consequence of Lemma 2.7 is that $T(x) \cap \mathcal{I}(\mathbf{z})$ contains at most one point.

Lemma 2.8. *For any $x \in \Gamma$, $T(x) \cap \mathcal{I}(\mathbf{z})$ contains at most one point for each $\mathbf{z} \in \mathbb{R}^m$.*

Proof. Take an arbitrary $x \in \Gamma$. Suppose there exist $\mathbf{x}_1 \neq \mathbf{x}_2$ such that $\mathbf{x}_1, \mathbf{x}_2 \in T(x) \cap \mathcal{I}(\hat{\mathbf{z}})$ for some $\hat{\mathbf{z}} \in \mathbb{R}^m$. Then $\mathbf{x}_3 = \mathbf{x}_2 - \mathbf{x}_1 \neq \mathbf{0}$, and, by Lemma 2.7, $\mathbf{x}_3 \in T(x)$. On the other hand,

$$\|\Pi_\perp(\mathbf{x}_3)\| = \|\Pi_\perp(\mathbf{x}_2) - \Pi_\perp(\mathbf{x}_1)\| = \|\hat{\mathbf{z}} - \hat{\mathbf{z}}\| = 0.$$

Since $\mathbf{x}_3 \neq \mathbf{0}$, we have $\|\Pi(\mathbf{x}_3)\| > \|\Pi_\perp(\mathbf{x}_3)\| = 0$. Then $\mathbf{x}_3 \in \text{int}(\mathcal{C}(\mathbf{0}))$. This constitutes a contradiction. \square

In fact, $T(x) \cap \mathcal{I}(\mathbf{z})$ contains exactly one point for every $\mathbf{z} \in \mathbb{R}^m$.

Lemma 2.9. *For any $x \in \Gamma$, $T(x) \cap \mathcal{I}(\mathbf{z})$ contains exactly one point for each $\mathbf{z} \in \mathbb{R}^m$.*

Proof. Take an arbitrary $x \in \Gamma$. By Lemma 2.8, we only need to prove that $T(x) \cap \mathcal{I}(\mathbf{z}) \neq \emptyset$ for any $\mathbf{z} \in \mathbb{R}^m$. We will use the Ważewski theorem (see Appendix A), which requires us to work with a *flow*. Thus, we append $\dot{\tau} = 1$ to (2.8) with the argument t in $\Phi(t, x)$ replaced by τ to form an autonomous system

$$\begin{aligned} \dot{\tau} &= 1, \\ \dot{\mathbf{x}} &= DF(\Phi(\tau, x)) \mathbf{x}, \end{aligned}$$

where $(\tau, \mathbf{x}) \in (-\epsilon_x, \infty) \times X$ with $\epsilon_x > 0$ being a constant that depends on the chosen $x \in \Gamma$. Let $\Psi_x(t, (\tau, \mathbf{x}))$ be the flow generated by the above system:

$$\Psi_x(t, (\tau, \mathbf{x})) = (t + \tau, Q(t, \Phi(\tau, x)) \mathbf{x}).$$

We define a set $\mathcal{W} \subset (-\epsilon_x, \infty) \times X$ as follows:

$$\begin{aligned} \mathcal{W} &:= \{(\tau, \mathbf{x}) : \tau \geq 0, \mathbf{x} \notin \text{int}(\mathcal{C}(\mathbf{0}))\} \\ (2.20) \quad &= \{(\tau, \mathbf{x}) : \tau \geq 0, \mathcal{L}(\mathbf{x}, \mathbf{0}) \leq 0\}. \end{aligned}$$

Since \mathcal{W} is a closed subset of $(-\epsilon_x, \infty) \times X$, it automatically satisfies the condition (W1) in the definition of a Ważewski set (see Appendix A), i.e., if $(\tau, \mathbf{x}) \in \mathcal{W}$ and

$\Psi_x([0, t], (\tau, \mathbf{x})) \subset \text{cl}(\mathcal{W})$ then $\Psi_x([0, t], (\tau, \mathbf{x})) \subset \mathcal{W}$. Next, we define the sets \mathcal{W}^0 and \mathcal{W}^- as follows:

$$\begin{aligned}\mathcal{W}^0 &:= \{(\tau, \mathbf{x}) \in \mathcal{W} : \exists t > 0 \text{ such that } \Psi_x(t, (\tau, \mathbf{x})) \notin \mathcal{W}\}, \\ \mathcal{W}^- &:= \{(\tau, \mathbf{x}) \in \mathcal{W} : \Psi_x([0, t], (\tau, \mathbf{x})) \not\subset \mathcal{W} \text{ for all } t > 0\}.\end{aligned}$$

Note that \mathcal{W}^0 is the set of points that do not stay in \mathcal{W} forever under the flow Ψ_x in forward time, and \mathcal{W}^- is the set of points that immediately leave \mathcal{W} in forward time. Clearly, $\mathcal{W}^- \subseteq \mathcal{W}^0 \subseteq \mathcal{W}$. In order to verify that \mathcal{W} is a Ważewski set, we need to check the condition (W2), that is, \mathcal{W}^- is closed relative to \mathcal{W}^0 .

It is obvious that $\mathcal{W}^- \subseteq \partial\mathcal{W}$, which is the boundary of \mathcal{W} and is the union of two disjoint sets $S_1 := \{(\tau, \mathbf{x}) : \tau \geq 0, \mathcal{L}(\mathbf{x}, \mathbf{0}) = 0\}$ and $S_2 := \{(0, \mathbf{x}) : \mathcal{L}(\mathbf{x}, \mathbf{0}) < 0\}$. For any $(\tau, \mathbf{x}) \in S_1$ with $\mathbf{x} \neq \mathbf{0}$, by the fact that $\Phi(\tau, x) \in \Gamma \subset U$, (2.15) holds, and it leads to

$$\left[\frac{d}{dt} \mathcal{L}(Q(t, \Phi(\tau, x)) \mathbf{x}, \mathbf{0}) \right]_{t=0} > 0.$$

Thus, $(\tau, \mathbf{x}) \in \mathcal{W}^-$ if $(\tau, \mathbf{x}) \in S_1$ and $\mathbf{x} \neq \mathbf{0}$. For each $(0, \mathbf{x}) \in S_2$, by the continuity of \mathcal{L} , there exists a $t_0 > 0$ such that $\mathcal{L}(Q(t, \Phi(0, x)) \mathbf{x}, \mathbf{0}) < 0$ for all $t \in [0, t_0]$. It follows that $\Psi_x(t, (0, \mathbf{x})) = (t, Q(t, \Phi(0, x)) \mathbf{x}) \in \mathcal{W}$ for all $t \in [0, t_0]$. Thus, $\mathcal{W}^- \cap S_2 = \emptyset$. Furthermore, note that for any $(\tau, \mathbf{0}) \in S_1$, $\Psi_x(t, (\tau, \mathbf{0})) = (t + \tau, \mathbf{0}) \in S_1 \subset \mathcal{W}$ for all $t \geq 0$. Altogether, we have simultaneously

$$(2.21) \quad \mathcal{W}^0 \cap \{(\tau, \mathbf{0}) : \tau \geq 0\} = \emptyset,$$

$$(2.22) \quad \mathcal{W}^- = S_1 \setminus \{(\tau, \mathbf{0}) : \tau \geq 0\} = \{(\tau, \mathbf{x}) : \tau \geq 0, \mathbf{x} \neq \mathbf{0}, \mathcal{L}(\mathbf{x}, \mathbf{0}) = 0\}.$$

In addition, by (2.21) and (2.22), we have

$$\mathcal{W}^0 \setminus \mathcal{W}^- \subseteq \{(\tau, \mathbf{x}) : \tau \geq 0, \mathcal{L}(\mathbf{x}, \mathbf{0}) < 0\}.$$

Thus, for every $(\tau, \mathbf{x}) \in \mathcal{W}^0 \setminus \mathcal{W}^-$, there is a neighborhood \mathcal{V} , which contains (τ, \mathbf{x}) and is open in $\mathbb{R} \times X$, such that $\mathcal{V} \cap \mathcal{W}^- = \emptyset$. Then, $\mathcal{W}^0 \setminus \mathcal{W}^-$ is open relative to \mathcal{W}^0 , and \mathcal{W}^- is closed relative to \mathcal{W}^0 . Therefore, \mathcal{W} is a Ważewski set. By the Ważewski theorem, there exists a continuous function $\mathcal{R} : \mathcal{W}^0 \times [0, 1] \rightarrow \mathcal{W}^0$ such that \mathcal{R} is a strong deformation retraction of \mathcal{W}^0 onto \mathcal{W}^- .

We now prove that $T(x) \cap \mathcal{I}(\mathbf{z}) \neq \emptyset$ for any $\mathbf{z} \in \mathbb{R}^m$. Since $\mathbf{0} \in T(x)$, we only need to show that $T(x) \cap \mathcal{I}(\mathbf{z}) \neq \emptyset$ for any $\mathbf{z} \in \mathbb{R}^m$ with $\|\mathbf{z}\| > 0$. We prove this by contradiction.

Assume that $T(x) \cap \mathcal{I}(\hat{\mathbf{z}}) = \emptyset$ for some $\hat{\mathbf{z}} \in \mathbb{R}^m$ with $\|\hat{\mathbf{z}}\| > 0$. Let $\hat{R} := \|\hat{\mathbf{z}}\|$. Then by the definition of $T(x)$ (see (2.12)) and (2.20), we have

$$\{(0, (\mathbf{a}, \hat{\mathbf{z}})) : \|\mathbf{a}\| \leq \hat{R}\} = \{(0, \mathbf{x}) : \mathbf{x} \in \mathcal{I}(\hat{\mathbf{z}})\} \cap \mathcal{W} \subset \mathcal{W}^0.$$

Thus the set $\{(0, (\mathbf{a}, \hat{\mathbf{z}})) : \|\mathbf{a}\| \leq \hat{R}\}$ is contained inside the domain of the continuous function $\mathcal{R}(\cdot, 1) : \mathcal{W}^0 \rightarrow \mathcal{W}^-$.

Moreover, for all $(\tau, \mathbf{x}) \in \mathcal{W}^-$, $\|\Pi(\mathbf{x})\| = \|\Pi_\perp(\mathbf{x})\| > 0$ according to (2.22). Thus we can define a projection $\mathcal{P} : \mathcal{W}^- \rightarrow \{\mathbf{a} \in \mathbb{R}^n : \|\mathbf{a}\| = \hat{R}\}$ as follows:

$$\mathcal{P}(\tau, \mathbf{x}) := \hat{R} \frac{\Pi(\mathbf{x})}{\|\Pi(\mathbf{x})\|}.$$

By taking the composition of \mathcal{P} , $\mathcal{R}(\cdot, 1)$, and the map $\mathbf{a} \mapsto (0, (\mathbf{a}, \hat{\mathbf{z}}))$, we obtain a continuous map $\mathcal{G} : \{\mathbf{a} \in \mathbb{R}^n : \|\mathbf{a}\| \leq \hat{R}\} \rightarrow \{\mathbf{a} \in \mathbb{R}^n : \|\mathbf{a}\| = \hat{R}\}$ as follows:

$$\mathcal{G}(\mathbf{a}) := \mathcal{P} \circ \mathcal{R}((0, (\mathbf{a}, \hat{\mathbf{z}})), 1).$$

Since \mathcal{R} is a strong deformation retraction of \mathcal{W}^0 onto \mathcal{W}^- , $\mathcal{R}((\tau, \mathbf{x}), 1) = (\tau, \mathbf{x})$ for all $(\tau, \mathbf{x}) \in \mathcal{W}^-$. Note that for all \mathbf{a} with $\|\mathbf{a}\| = \hat{R}$, $(0, (\mathbf{a}, \hat{\mathbf{z}})) \in \mathcal{W}^-$ according to (2.22). Thus, for all \mathbf{a} with $\|\mathbf{a}\| = \hat{R}$,

$$\begin{aligned} \mathcal{G}(\mathbf{a}) &= \mathcal{P} \circ \mathcal{R}((0, (\mathbf{a}, \hat{\mathbf{z}})), 1) \\ &= \mathcal{P}(0, (\mathbf{a}, \hat{\mathbf{z}})) \\ &= \mathbf{a}. \end{aligned}$$

The existence of such a \mathcal{G} contradicts the fact that there is no retraction that maps a closed n -ball onto its boundary (i.e., an $(n-1)$ -sphere). \square

Based on Lemma 2.7 and Lemma 2.9, we can associate each $x \in \Gamma$ with a unique linear operator $H(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ such that $T(x)$ is the graph of $H(x)$, i.e.,

$$(2.23) \quad T(x) = \{(H(x)\mathbf{z}, \mathbf{z}) : \mathbf{z} \in \mathbb{R}^m\}.$$

In addition, it follows from (2.22) that for each fixed $x \in \Gamma$,

$$\mathcal{L}((H(x)\mathbf{z}, \mathbf{z}), \mathbf{0}) < 0 \text{ for all } \mathbf{z} \in \mathbb{R}^m \text{ with } \|\mathbf{z}\| = 1.$$

Then the compactness of the set $\{\mathbf{z} \in \mathbb{R}^m : \|\mathbf{z}\| = 1\}$ implies that for any $x \in \Gamma$,

$$(2.24) \quad \|H(x)\| = \max_{\|\mathbf{z}\|=1} \|H(x)\mathbf{z}\| < 1.$$

2.2.2. *The map $x \mapsto H(x)$ is continuous.* We first prove the following lemma.

Lemma 2.10. *For any fixed $\mathbf{z} \in \mathbb{R}^m$, the map from Γ into \mathbb{R}^n : $x \mapsto H(x)\mathbf{z}$ is continuous.*

Proof. Take an arbitrary $\hat{x} \in \Gamma$ and any $\mathbf{z} \in \mathbb{R}^m$. We prove that $\lim_{x \rightarrow \hat{x}} H(x)\mathbf{z} = H(\hat{x})\mathbf{z}$ by contradiction. Suppose that there exists a constant $\varepsilon > 0$ and a sequence $\{x_1, x_2, x_3, \dots\} \subset \Gamma$ such that $x_i \rightarrow \hat{x}$ and at the same time

$$(2.25) \quad \|H(x_i)\mathbf{z} - H(\hat{x})\mathbf{z}\| > \varepsilon$$

for all $i = 1, 2, 3, \dots$. Since $\|H(x_i)\mathbf{z}\| < \|\mathbf{z}\|$ for $i = 1, 2, 3, \dots$, we can take a convergent subsequence $\{H(x_{i_1})\mathbf{z}, H(x_{i_2})\mathbf{z}, H(x_{i_3})\mathbf{z}, \dots\}$ and denote its limit by $\hat{\mathbf{a}}$, i.e., $H(x_{i_k})\mathbf{z} \rightarrow \hat{\mathbf{a}}$. Then (2.25) implies that $\|\hat{\mathbf{a}} - H(\hat{x})\mathbf{z}\| \geq \varepsilon$. Thus $\hat{\mathbf{x}} := (\hat{\mathbf{a}}, \mathbf{z}) \notin T(\hat{x})$, and there exists a $t_0 \geq 0$ such that $Q(t_0, \hat{x})\hat{\mathbf{x}} \in \text{int}(\mathcal{C}(\mathbf{0}))$. Let q be the radius of a closed ball that is centered at $Q(t_0, \hat{x})\hat{\mathbf{x}}$ and contained in $\text{int}(\mathcal{C}(\mathbf{0}))$.

Let $\mathbf{x}_{i_k} := (H(x_{i_k})\mathbf{z}, \mathbf{z}) \in T(x_{i_k})$ for $k = 1, 2, 3, \dots$. Consider the convergent sequence $\{(x_{i_1}, \mathbf{x}_{i_1}), (x_{i_2}, \mathbf{x}_{i_2}), (x_{i_3}, \mathbf{x}_{i_3}), \dots\} \subset \Gamma \times X$ with limit $(\hat{x}, \hat{\mathbf{x}})$. Since $Q(t_0, x)\mathbf{x}$ depends on (x, \mathbf{x}) continuously for all $(x, \mathbf{x}) \in \Gamma \times X$, we can choose an i_{k^*} large enough such that

$$\|Q(t_0, x_{i_{k^*}})\mathbf{x}_{i_{k^*}} - Q(t_0, \hat{x})\hat{\mathbf{x}}\| \leq q.$$

Then $Q(t_0, x_{i_{k^*}})\mathbf{x}_{i_{k^*}} \in \text{int}(\mathcal{C}(\mathbf{0}))$. However, this is impossible since $\mathbf{x}_{i_{k^*}} \in T(x_{i_{k^*}})$ by our construction. \square

Since $H(x)$ is a linear operator from \mathbb{R}^m to \mathbb{R}^n for each $x \in \Gamma$, Lemma 2.10 implies the continuity of the map $x \mapsto H(x)$ for all $x \in \Gamma$.

2.2.3. $H(h(z), z)$ is the derivative of h at z . The final step of establishing the C^1 smoothness of h is to prove the next lemma.

Lemma 2.11. $h : K \rightarrow \mathbb{R}^n$ is differentiable at all $z \in K$. In addition, $Dh(z) = H(h(z), z)$ for all $z \in K$.

Proof. We will show that $H(x)$ satisfies the definition of the derivative of h , i.e., for any $z \in K$ and correspondingly $x = (h(z), z) \in \Gamma$,

$$\lim_{\|z\| \rightarrow 0} \frac{1}{\|z\|} \|h(z + z) - h(z) - H(x)z\| = 0.$$

We prove this by contradiction. Assume that for some $\hat{z} \in K$ and correspondingly $\hat{x} = (h(\hat{z}), \hat{z}) \in \Gamma$, there exist $\{z_1, z_2, z_3, \dots\}$, a sequence of *non-zero* vectors in \mathbb{R}^m with $\|z_i\| \rightarrow 0$, and a constant $\varepsilon > 0$ such that $\hat{z} + z_i \in K$ and

$$(2.26) \quad \frac{1}{\|z_i\|} \|h(\hat{z} + z_i) - h(\hat{z}) - H(\hat{x})z_i\| > \varepsilon$$

for all $i = 1, 2, 3, \dots$. Let $\mathbf{a}_i = (h(\hat{z} + z_i) - h(\hat{z})) / \|z_i\|$ for $i = 1, 2, 3, \dots$. By Lemma 2.3, we have $\|\mathbf{a}_i\| < 1$ for any i . Let $\mathbf{x}_i = (\mathbf{a}_i, z_i / \|z_i\|)$ for $i = 1, 2, 3, \dots$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots\}$ is a bounded sequence in X . Take a convergent subsequence $\{\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \mathbf{x}_{i_3}, \dots\}$, and denote its limit by $\hat{\mathbf{x}} = (\hat{\mathbf{a}}, \hat{\mathbf{z}})$. Note that $\mathbf{a}_{i_k} \rightarrow \hat{\mathbf{a}}$ and $z_{i_k} / \|z_{i_k}\| \rightarrow \hat{\mathbf{z}}$. Then (2.26) implies that $\|\hat{\mathbf{a}} - H(\hat{x})\hat{\mathbf{z}}\| \geq \varepsilon$. Thus $\hat{\mathbf{x}} \notin T(\hat{x})$, and there exists a $t_0 \geq 0$ such that $Q(t_0, \hat{x})\hat{\mathbf{x}} \in \text{int}(\mathcal{C}(\mathbf{0}))$. Let q be the radius of a closed ball that is centered at $Q(t_0, \hat{x})\hat{\mathbf{x}}$ and contained in $\text{int}(\mathcal{C}(\mathbf{0}))$.

Furthermore, by the constructions of \mathbf{a}_i and \mathbf{x}_i , we have that for any i_k ,

$$\begin{aligned} \hat{x} + \|z_{i_k}\| \mathbf{x}_{i_k} &= (h(\hat{z}) + \|z_{i_k}\| \mathbf{a}_{i_k}, \hat{z} + z_{i_k}) \\ &= (h(\hat{z} + z_{i_k}), \hat{z} + z_{i_k}) \in \Gamma. \end{aligned}$$

Then $\Phi(t_0, \hat{x}) \in \Gamma$ and $\Phi(t_0, \hat{x} + \|z_{i_k}\| \mathbf{x}_{i_k}) \in \Gamma$ according to the positive invariance of Γ . Note that $\|z_{i_k}\| \mathbf{x}_{i_k} \neq \mathbf{0}$ for any i_k . Thus $\Phi(t_0, \hat{x} + \|z_{i_k}\| \mathbf{x}_{i_k}) \notin \mathcal{C}(\Phi(t_0, \hat{x}))$ (see the proof of Lemma 2.3) for any i_k .

On the other hand, note that $Q(t_0, \hat{x})$ is just the derivative of $\Phi(t_0, x)$ with respect to x evaluated at $x = \hat{x}$. Recall that by our construction, $\{\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \mathbf{x}_{i_3}, \dots\}$ is bounded and $\|z_{i_k}\| \rightarrow 0$. Then there exists an N such that for any $i_k \geq N$,

$$(2.27) \quad \begin{aligned} \frac{q}{3} &\geq \frac{1}{\|z_{i_k}\|} \|\Phi(t_0, \hat{x} + \|z_{i_k}\| \mathbf{x}_{i_k}) - \Phi(t_0, \hat{x}) - Q(t_0, \hat{x})(\|z_{i_k}\| \mathbf{x}_{i_k})\| \\ &= \left\| \frac{1}{\|z_{i_k}\|} (\Phi(t_0, \hat{x} + \|z_{i_k}\| \mathbf{x}_{i_k}) - \Phi(t_0, \hat{x})) - Q(t_0, \hat{x}) \mathbf{x}_{i_k} \right\|. \end{aligned}$$

In addition, we can choose an $i_{k^*} \geq N$ such that

$$(2.28) \quad \|Q(t_0, \hat{x}) \mathbf{x}_{i_{k^*}} - Q(t_0, \hat{x}) \hat{\mathbf{x}}\| \leq \frac{q}{3}.$$

Combining (2.27) and (2.28), we obtain

$$\left\| \frac{1}{\|z_{i_{k^*}}\|} (\Phi(t_0, \hat{x} + \|z_{i_{k^*}}\| \mathbf{x}_{i_{k^*}}) - \Phi(t_0, \hat{x})) - Q(t_0, \hat{x}) \hat{\mathbf{x}} \right\| \leq \frac{q}{3} + \frac{q}{3} < q,$$

which implies that

$$\frac{1}{\|z_{i_{k^*}}\|} (\Phi(t_0, \hat{x} + \|z_{i_{k^*}}\| \mathbf{x}_{i_{k^*}}) - \Phi(t_0, \hat{x})) \in \text{int}(\mathcal{C}(\mathbf{0})).$$

It follows that $\Phi(t_0, \hat{x} + \|z_{i_{k^*}}\| \mathbf{x}_{i_{k^*}}) \in \text{int}(\mathcal{C}(\Phi(t_0, \hat{x})))$. This constitutes a contradiction. \square

Note that the map from K onto Γ : $z \mapsto (h(z), z)$ is continuous. Thus $Dh(z) = H(h(z), z)$ is continuous with respect to z on K . Finally, it follows from (2.24) that $\|Dh(z)\| < 1$ for all $z \in K$. This concludes the proof of the C^1 smoothness of h .

3. PROOF OF THEOREM 1.2

We begin with the assumption that the positively invariant set Γ is a C^1 manifold embedded in $\mathbb{R}^n \times \mathbb{R}^m$ as the graph of the C^1 function h_0 from an open, convex $K_0 \subseteq \mathbb{R}^m$ to \mathbb{R}^n . The proof of the C^r smoothness of h_0 proceeds as follows. In Subsection 3.1, we derive a system that describes the dynamics of Dh_0 along trajectories in the C^1 manifold Γ . In Subsection 3.2, we show that this system can also be put into the form of (1.3). Then by applying Theorem 1.1, we prove that the set $\{(Dh_0(z), z) : z \in K_0\}$ can be smoothly embedded into an appropriate Euclidean space as a C^1 manifold, which implies the C^2 smoothness of h_0 . Finally, in Subsection 3.3, we demonstrate how to apply this argument inductively to establish the C^r smoothness of h_0 .

3.1. The Dynamics of Dh_0 along Trajectories in Γ . We shall continue to use some of the notations that have been introduced in Subsection 2.2. However, from now on we represent all linear operators as matrices. In particular, for each $x \in \Gamma$, the $(n+m) \times (n+m)$ matrix function $t \mapsto Q(t, x)$ is just the *fundamental matrix solution* of (2.8) with parameter x , satisfying the matrix differential equation

$$\dot{Q}(t, x) = DF(\Phi(t, x)) Q(t, x)$$

with $Q(0, x) = I_{n+m}$, which is the $(n+m) \times (n+m)$ identity matrix.

Restricting the z -component of (1.3) to Γ , we obtain

$$\dot{z} = g(h_0(z), z),$$

which generates a flow $\phi(t, z)$ on K_0 . Due to the positive invariance of Γ , $\phi(t, z)$ is related to $\Phi(t, x)$ in the following obvious way:

$$(h_0(\phi(t, z)), \phi(t, z)) = \Phi(t, (h_0(z), z)) \text{ for any } z \in K_0 \text{ and any } t \geq 0.$$

Then the dynamics of Dh_0 along trajectories in Γ are governed by a matrix Riccati differential equation as stated in the next lemma.

Lemma 3.1. *Consider any system in the form of (1.3) with f and g at least C^1 . If $h_0 : K_0 \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$ is C^1 and $\Gamma = \{(h_0(z), z) : z \in K_0\}$ is positively invariant, then for each fixed $z \in K_0$, the matrix function $Dh_0(\phi(\cdot, z)) : [0, \infty) \rightarrow \mathbb{R}^{n \times m}$ (the space of $n \times m$ real matrices) is a solution of the following matrix Riccati differential equation:*

$$(3.1) \quad \dot{V}^1 = D_a f V^1 - V^1 D_z g - V^1 D_a g V^1 + D_z f,$$

where $V^1 \in \mathbb{R}^{n \times m}$ and $D_a f$, $D_z f$, $D_a g$, and $D_z g$ are all evaluated along the trajectory $(h_0(\phi(t, z)), \phi(t, z))$ for $t \geq 0$.

Proof. Take an arbitrary $z \in K_0$ and correspondingly $x = (h_0(z), z) \in \Gamma$. First, we show that the matrix function $Dh_0(\phi(\cdot, z)) : [0, \infty) \rightarrow \mathbb{R}^{n \times m}$ is continuously differentiable.

Note that for any $t \geq 0$, $Q(t, x)$ maps the tangent space of Γ at x to the tangent space of Γ at $\Phi(t, x)$. It follows that for any $z \in \mathbb{R}^m$,

$$(3.2) \quad Q(t, x) \begin{pmatrix} Dh_0(z) \\ I_m \end{pmatrix} z = \begin{pmatrix} Dh_0(\phi(t, z)) \\ I_m \end{pmatrix} (0_{m \times n} \quad I_m) Q(t, x) \begin{pmatrix} Dh_0(z) \\ I_m \end{pmatrix} z.$$

Note that for any $\mathbf{z} \in \mathbb{R}^m$ with $\|\mathbf{z}\| > 0$, the left hand side of (3.2) represents a solution (of (2.8) with the parameter $x = (h_0(z), z) \in \Gamma$) that cannot reach the origin for any finite $t \geq 0$. Thus it must be true that for any $t \geq 0$, the null space of the product of the second, third, and forth matrices on the right-hand side of (3.2) is the trivial subspace of \mathbb{R}^m . Therefore, this product produces an invertible matrix for any $t \geq 0$. Solving (3.2) for $Dh_0(\phi(t, z))$, we obtain

$$Dh_0(\phi(t, z)) = (I_n \ 0_{n \times m}) Q(t, x) \begin{pmatrix} Dh_0(z) \\ I_m \end{pmatrix} \left((0_{m \times n} \ I_m) Q(t, x) \begin{pmatrix} Dh_0(z) \\ I_m \end{pmatrix} \right)^{-1}.$$

In view of the differentiability of $Q(t, x)$ with respect to t , $Dh_0(\phi(t, z))$ is continuously differentiable with respect to t for all $t \geq 0$.

Next, we derive a differential equation for $Dh_0(\phi(t, z))$. Choose any $t \geq 0$, and consider (2.14) at $(Dh_0(\phi(t, z)) \mathbf{z}, \mathbf{z}) \in T_{\Phi(t, x)}\Gamma$ for an arbitrary $\mathbf{z} \in \mathbb{R}^m$. Then we have

$$\begin{aligned} \left(\frac{d}{dt} Dh_0(\phi(t, z)) \right) \mathbf{z} + Dh_0(\phi(t, z)) \dot{\mathbf{z}} \\ = D_a f(\Phi(t, x)) Dh_0(\phi(t, z)) \mathbf{z} + D_z f(\Phi(t, x)) \mathbf{z}, \\ \dot{\mathbf{z}} = D_a g(\Phi(t, x)) Dh_0(\phi(t, z)) \mathbf{z} + D_z g(\Phi(t, x)) \mathbf{z}. \end{aligned}$$

Combining these two equations and reorganizing terms, we obtain that $Dh_0(\phi(t, z))$ satisfies the matrix Riccati differential equation (3.1) with the coefficient matrices $D_a f$, $D_z f$, $D_a g$, and $D_z g$ all evaluated along the trajectory $\Phi(t, x) = (h_0(\phi(t, z)), \phi(t, z))$ for $t \geq 0$. \square

3.2. The C^2 Smoothness of h_0 . For each $z \in K_0$, (3.1) defines a time varying system in $\mathbb{R}^{n \times m}$ along the trajectory $(h_0(\phi(t, z)), \phi(t, z))$ for $t \geq 0$. Thus, we can couple (3.1) with the underlying equation $\dot{z} = g(h_0(z), z)$, which generates the flow $\phi(t, z)$, to form a system in $\mathbb{R}^{n \times m} \times K_0$ as follows:

(3.3a)

$$\dot{V}^1 = D_a f(h_0(z), z) V^1 - V^1 D_z g(h_0(z), z) - V^1 D_a g(h_0(z), z) V^1 + D_z f(h_0(z), z),$$

(3.3b)

$$\dot{z} = g(h_0(z), z),$$

where $V^1 \in \mathbb{R}^{n \times m}$ and $z \in K_0$.

In the subsequent analysis, it is more convenient to consider the “vectorization” of the matrix differential equation (3.3a). Recall that the *vectorization* of a matrix is a transformation that converts the matrix into a column vector by stacking the columns of the matrix. Let $\text{vec}_{n,m} : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{nm}$ be the vectorization of $n \times m$ matrices, and let $\text{vec}_{n,m}^{-1} : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{n \times m}$ be the inverse of $\text{vec}_{n,m}$. Notice that for any $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$, and $P \in \mathbb{R}^{n \times m}$,

$$\text{vec}_{n,m}(APB) = (B^T \otimes A) \text{vec}_{n,m}(P),$$

where “ \otimes ” denotes the Kronecker (or tensor) product of two matrices. Then (3.3a) can be converted into

$$\begin{aligned} \dot{v}^1 = & \left(I_m \otimes D_a f(h_0(z), z) - (D_z g(h_0(z), z))^T \otimes I_n \right) v^1 \\ & - \left(I_m \otimes (\text{vec}_{n,m}^{-1}(v^1) D_a g(h_0(z), z)) \right) v^1 + \text{vec}_{n,m}(D_z f(h_0(z), z)), \end{aligned}$$

where $v^1 := \text{vec}_{n,m}(V^1) \in \mathbb{R}^{nm}$.

Furthermore, we rescale z to $\zeta^1 := \sigma_1^{-1}z$ with the scaling factor $\sigma_1 > 0$ to be determined later. Let $K_1 := \{\zeta^1 \in \mathbb{R}^m : \sigma_1 \zeta^1 \in K_0\}$. Then we transform (3.3) into a system defined on $\mathbb{R}^{nm} \times K_1$ as follows:

$$(3.4) \quad \begin{aligned} \dot{v}^1 &= f_1(v^1, \zeta^1), \\ \dot{\zeta}^1 &= \gamma_1(\zeta^1), \end{aligned}$$

where $v^1 \in \mathbb{R}^{nm}$, $\zeta^1 \in K_1$, and the functions $\gamma_1 : K_1 \rightarrow \mathbb{R}^m$ and $f_1 : \mathbb{R}^{nm} \times K_1 \rightarrow \mathbb{R}^{nm}$ are defined as follows:

$$(3.5) \quad \gamma_1(\zeta^1) := \frac{1}{\sigma_1} g(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1),$$

$$(3.6) \quad \begin{aligned} f_1(v^1, \zeta^1) &:= \left(I_m \otimes D_a f(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1) - (D_z g(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1))^T \otimes I_n \right) v^1 \\ &\quad - \left(I_m \otimes (\text{vec}_{n,m}^{-1}(v^1) D_a g(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1)) \right) v^1 \\ &\quad + \text{vec}_{n,m}(D_z f(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1)). \end{aligned}$$

Choose J_1 , a bounded, open subset of \mathbb{R}^{nm} , as follows:

$$J_1 := \{v^1 \in \mathbb{R}^{nm} : \|\text{vec}_{n,m}^{-1}(v^1)\| < \eta\},$$

where, as postulated in Hypothesis 4, η is a positive constant such that $\|Dh_0(z)\| < \eta$ for all $z \in K_0$. Obviously $J_1 \times K_1$ satisfies Hypothesis 1. In what follows, we will verify that the restriction of (3.4) to $J_1 \times K_1$ satisfies Hypothesis 2.

Lemma 3.2. *f_1 is C^1 on $J_1 \times K_1$, and γ_1 is C^1 on K_1 . Furthermore, there exists a constant $L_1 > 0$ independent of the choice of σ_1 such that $\|D_{\zeta^1} f_1(v^1, \zeta^1)\| \leq \sigma_1 L_1$ for any $(v^1, \zeta^1) \in J_1 \times K_1$.*

Proof. By Hypothesis 5, f and g are C^r ($r \geq 2$) with their first to r -th derivatives all bounded on U_0 . In addition, h_0 is C^1 with $\|Dh_0(z)\| < \eta$ for all $z \in K_0$. Then by inspecting (3.5) and (3.6), we obtain the abovementioned properties of f_1 and γ_1 . \square

Define a continuous function $\alpha_1 : K_1 \rightarrow \mathbb{R}$ as follows:

$$(3.7) \quad \alpha_1(\zeta^1) := \alpha(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1) - \ell(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1) - 2\eta \|D_a g(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1)\|.$$

By (1.7c) of Hypothesis 5, we have that $\alpha_1(\zeta^1) > 0$ for all $\zeta^1 \in K_1$.

Lemma 3.3. *For any $(v_0^1, \zeta_0^1) \in J_1 \times K_1$ and any $v^1 \in \mathbb{R}^{nm}$,*

$$(3.8) \quad \langle v^1, D_{v^1} f_1(v_0^1, \zeta_0^1) v^1 \rangle \geq \alpha_1(\zeta_0^1) \|v^1\|^2.$$

Proof. Take an arbitrary $(v_0^1, \zeta_0^1) \in J_1 \times K_1$. It follows from (3.6) that

$$\begin{aligned} D_{v^1} f_1(v_0^1, \zeta_0^1) &= I_m \otimes D_a f(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1) - (D_z g(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1))^T \otimes I_n \\ &\quad - I_m \otimes (\text{vec}_{n,m}^{-1}(v_0^1) D_a g(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1)) \\ &\quad - (D_a g(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1) \text{vec}_{n,m}^{-1}(v_0^1))^T \otimes I_n. \end{aligned}$$

Taking account of (1.7a), it is straightforward to verify that for all $v^1 \in \mathbb{R}^{nm}$,

$$\langle v^1, (I_m \otimes D_a f) v^1 \rangle \geq \alpha \|v^1\|^2,$$

where both $D_a f$ and α are evaluated at $(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1)$.

Next, take any $v^1 \in \mathbb{R}^{nm}$ and correspondingly $V^1 = \text{vec}_{n,m}^{-1}(v^1)$. We have that

$$\langle v^1, ((D_z g)^T \otimes I_n) v^1 \rangle = \langle \text{vec}_{n,m}(V^1), \text{vec}_{n,m}(V^1 D_z g) \rangle$$

$$\begin{aligned}
&= \langle \text{vec}_{m,n}(V^{1T}), \text{vec}_{m,n}((D_z g)^T V^{1T}) \rangle \\
&= \langle \text{vec}_{m,n}(V^{1T}), (I_n \otimes (D_z g)^T) \text{vec}_{m,n}(V^{1T}) \rangle \\
&= \langle \text{vec}_{m,n}(V^{1T}), (I_n \otimes (D_z g)^T)^T \text{vec}_{m,n}(V^{1T}) \rangle \\
&= \langle \text{vec}_{m,n}(V^{1T}), (I_n \otimes D_z g) \text{vec}_{m,n}(V^{1T}) \rangle \\
&\leq \ell \| \text{vec}_{m,n}(V^{1T}) \|^2 = \ell \| v^1 \|^2,
\end{aligned}$$

where both $D_z g$ and ℓ are evaluated at $(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1)$ throughout and the inequality follows from (1.7b).

Similarly, we obtain that for any $v^1 \in \mathbb{R}^{nm}$,

$$\begin{aligned}
\left| \langle v^1, (I_m \otimes (\text{vec}_{n,m}^{-1}(v_0^1) D_a g)) v^1 \rangle \right| &\leq \| \text{vec}_{n,m}^{-1}(v_0^1) D_a g \| \| v^1 \|^2 < \eta \| D_a g \| \| v^1 \|^2, \\
\left| \langle v^1, ((D_a g \text{vec}_{n,m}^{-1}(v_0^1))^T \otimes I_n) v^1 \rangle \right| &\leq \| D_a g \text{vec}_{n,m}^{-1}(v_0^1) \| \| v^1 \|^2 < \eta \| D_a g \| \| v^1 \|^2,
\end{aligned}$$

where $D_a g$ is evaluated at $(h_0(\sigma_1 \zeta_0^1), \sigma_1 \zeta_0^1)$ throughout and $\| \text{vec}_{n,m}^{-1}(v_0^1) \| < \eta$ by our choice of J_1 .

Combining all the estimates above, we attain the inequality (3.8). \square

Define a nonnegative, continuous function $\ell_1 : K_1 \rightarrow \mathbb{R}$ as follows:

$$(3.9) \quad \ell_1(\zeta^1) := \eta \| D_a g(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1) \| + \ell(h_0(\sigma_1 \zeta^1), \sigma_1 \zeta^1).$$

Then in view of (3.5), the next lemma is a trivial consequence of (1.7b).

Lemma 3.4. *For any $\zeta_0^1 \in K_1$ and any $\zeta^1 \in \mathbb{R}^m$,*

$$(3.10) \quad \langle \zeta^1, D\gamma_1(\zeta_0^1) \zeta^1 \rangle \leq \ell_1(\zeta_0^1) \| \zeta^1 \|^2.$$

In addition, we have the estimate described in the following lemma.

Lemma 3.5. *There exists a sufficiently small σ_1 such that for any $\zeta^1 \in K_1$,*

$$(3.11) \quad \alpha_1(\zeta^1) \geq \ell_1(\zeta^1) + \sigma_1 L_1 + \frac{1}{2} c_r.$$

Proof. It follows from (1.7c) of Hypothesis 5 that for any $z \in K_0$,

$$(3.12) \quad \alpha(h_0(z), z) \geq 2\ell(h_0(z), z) + 3\eta \| D_a g(h_0(z), z) \| + c_r.$$

In addition, since $L_1 > 0$ is independent of the choice of σ_1 (see Lemma 3.2), we can choose σ_1 sufficiently small so that $\sigma_1 L_1 \leq \frac{1}{2} c_r$. Then by combining this with (3.12), we obtain the inequality (3.11) after replacing z with $\sigma_1 \zeta^1$ and rearranging terms according to (3.7) and (3.9). \square

Note that the ζ^1 -component of (3.4) is independent of v^1 . Thus Lemmas 3.2–3.5 verify that the restriction of (3.4) to $J_1 \times K_1$ satisfies Hypotheses 2.

Recall that $\| Dh_0(z) \| < \eta$ for all $z \in K_0$. Define $h_1 : K_1 \rightarrow J_1$ as follows:

$$(3.13) \quad h_1(\zeta^1) := \text{vec}_{n,m}(Dh_0(\sigma_1 \zeta^1)).$$

Now consider the set

$$\Gamma_1 := \{ (h_1(\zeta^1), \zeta^1) : \zeta^1 \in K_1 \} \subset J_1 \times K_1.$$

Clearly, Γ_1 is positively invariant under the flow of (3.4). Then, by applying Theorem 1.1, we obtain the following result.

Lemma 3.6. *Γ_1 is a C^1 manifold. In particular, $h_1 : K_1 \rightarrow J_1$ is C^1 with $\| Dh_1(\zeta^1) \| < 1$ for all $\zeta^1 \in K_1$.*

Note that $Dh_0(z) = \text{vec}_{n,m}^{-1}(h_1(\sigma_1^{-1}z))$ for any $z \in K_0$. Therefore, we have proven that $h_0 : K_0 \rightarrow \mathbb{R}^n$ is C^2 with its first and second derivatives bounded on K_0 . Furthermore,

$$D^2h_0(z)(z_1, z_2) = \sigma_1^{-1} \text{vec}_{n,m}^{-1}(Dh_1(\sigma_1^{-1}z) z_1) z_2$$

for any $z \in K_0$ and any $z_1, z_2 \in \mathbb{R}^m$.

3.3. The C^r Smoothness of h_0 . In this subsection, we apply the argument used in the previous subsection inductively to establish the C^r smoothness of h_0 for the case with $r \geq 3$.

We illustrate how the argument works again in the proof of the C^3 smoothness of h_0 . First, let $\phi_1(t, \zeta^1)$ be the flow on K_1 generated by $\dot{\zeta}^1 = \gamma_1(\zeta^1)$. Then by Lemma 3.1, we have that for each fixed $\zeta^1 \in K_1$, the matrix function $Dh_1(\phi_1(\cdot, \zeta^1)) : [0, \infty) \rightarrow \mathbb{R}^{nm \times m}$ is a solution of the matrix Riccati differential equation:

$$(3.14) \quad \dot{V}^2 = D_{v^1}f_1 V^2 - V^2 D\gamma_1 + D_{\zeta^1}f_1,$$

where $V^2 \in \mathbb{R}^{nm \times m}$ and $D_{v^1}f_1$, $D_{\zeta^1}f_1$, and $D\gamma_1$ are all evaluated along the trajectory $(h_1(\phi_1(t, \zeta^1)), \phi_1(t, \zeta^1))$ for $t \geq 0$. Compared to (3.1), the right-hand side of (3.14) does not include the term quadratic in V^2 since $D_{v^1}\gamma_1 = 0$. Next, we couple (3.14) with the underlying equation $\dot{\zeta}^1 = \gamma_1(\zeta^1)$ to form a system in $\mathbb{R}^{nm \times m} \times K_1$ as follows:

$$(3.15a) \quad \dot{V}^2 = D_{v^1}f_1(h_1(\zeta^1), \zeta^1) V^2 - V^2 D\gamma_1(\zeta^1) + D_{\zeta^1}f_1(h_1(\zeta^1), \zeta^1),$$

$$(3.15b) \quad \dot{\zeta}^1 = \gamma_1(\zeta^1).$$

We shall again consider the vectorization of the matrix differential equation (3.15a). Let $\text{vec}_{nm,m} : \mathbb{R}^{nm \times m} \rightarrow \mathbb{R}^{nm^2}$ be the vectorization of $nm \times m$ matrices, and let $\text{vec}_{nm,m}^{-1}$ be the inverse of $\text{vec}_{nm,m}$. Furthermore, we rescale ζ^1 to $\zeta^2 := \sigma_2^{-1}\zeta^1$ with the scaling factor $\sigma_2 > 0$ to be determined. Take $K_2 := \{\zeta^2 \in \mathbb{R}^m : \sigma_2\zeta^2 \in K_1\}$. Then we transform (3.15) into a system defined on $\mathbb{R}^{nm^2} \times K_2$ as follows:

$$(3.16) \quad \begin{aligned} \dot{v}^2 &= f_2(v^2, \zeta^2), \\ \dot{\zeta}^2 &= \gamma_2(\zeta^2), \end{aligned}$$

where $v^2 := \text{vec}_{nm,m}(V^2) \in \mathbb{R}^{nm^2}$, $\zeta^2 \in K_2$, and the functions $\gamma_2 : K_2 \rightarrow \mathbb{R}^m$ and $f_2 : \mathbb{R}^{nm^2} \times K_2 \rightarrow \mathbb{R}^{nm^2}$ are defined as follows:

$$(3.17) \quad \gamma_2(\zeta^2) := \frac{1}{\sigma_2} \gamma_1(\sigma_2 \zeta^2),$$

$$(3.18) \quad \begin{aligned} f_2(v^2, \zeta^2) &:= (I_m \otimes D_{v^1}f_1(h_1(\sigma_2 \zeta^2), \sigma_2 \zeta^2) - (D\gamma_1(\sigma_2 \zeta^2))^T \otimes I_{nm}) v^2 \\ &\quad + \text{vec}_{nm,m}(D_{\zeta^1}f_1(h_1(\sigma_2 \zeta^2), \sigma_2 \zeta^2)). \end{aligned}$$

Choose J_2 , a bounded, open subset of \mathbb{R}^{nm^2} , as follows:

$$J_2 := \{v^2 \in \mathbb{R}^{nm^2} : \|\text{vec}_{nm,m}^{-1}(v^2)\| < 1\}.$$

Obviously $J_2 \times K_2$ satisfies Hypothesis 1. In addition, the restriction of (3.16) to $J_2 \times K_2$ again satisfies Hypothesis 2 as verified below.

Lemma 3.7. *f_2 is C^1 on $J_2 \times K_2$, and γ_2 is C^2 on K_2 . Furthermore, there exists a constant $L_2 > 0$ independent of the choice of σ_2 such that $\|D_{\zeta^2}f_2(v^2, \zeta^2)\| \leq \sigma_2 L_2$ for any $(v^2, \zeta^2) \in J_2 \times K_2$.*

Proof. So far, we have proven that 1) h_0 is C^2 with Dh_0 and D^2h_0 both bounded on K_0 and 2) h_1 defined by (3.13) is C^1 with Dh_1 bounded on K_1 . Thus, by our assumption that f and g are C^r ($r \geq 3$ in this case) with their first to r -th derivatives all bounded on U_0 , now f_1 and γ_1 defined by (3.6) and (3.5) are both C^2 with their first and second derivatives bounded on $J_1 \times K_1$ and K_1 , respectively. In turn, f_2 is C^1 with $\|D_{\zeta^2} f_2\| \leq \sigma_2 L_2$ on $J_2 \times K_2$ for some $L_2 > 0$, which is independent of the choice of σ_2 . In addition, γ_2 is C^2 since it has the same smoothness as γ_1 . \square

Define continuous functions $\alpha_2 : K_2 \rightarrow \mathbb{R}$ and $\ell_2 : K_2 \rightarrow \mathbb{R}$ as follows:

$$(3.19) \quad \alpha_2(\zeta^2) := \alpha_1(\sigma_2 \zeta^2) - \ell_1(\sigma_2 \zeta^2),$$

$$(3.20) \quad \ell_2(\zeta^2) := \ell_1(\sigma_2 \zeta^2).$$

Clearly, α_2 is positive because of (3.11), and ℓ_2 is nonnegative as ℓ_1 is. In addition, we have a set of estimates for α_2 and ℓ_2 similar to what we have for α_1 and ℓ_1 in Lemmas 3.3–3.5.

Lemma 3.8. *For any $(v_0^2, \zeta_0^2) \in J_2 \times K_2$ and any $v^2 \in \mathbb{R}^{nm^2}$,*

$$(3.21) \quad \langle v^2, D_{v^2} f_2(v_0^2, \zeta_0^2) v^2 \rangle \geq \alpha_2(\zeta_0^2) \|v^2\|^2.$$

Proof. Consider an arbitrary $(v_0^2, \zeta_0^2) \in J_2 \times K_2$. Notice that $D_{v^2} f_2$ is simply the coefficient matrix left-multiplied to v^2 in (3.18). By the same arguments as what we have used in the proof of Lemma 3.3 and incorporating the estimates (3.8) and (3.10), we have that for all $v^2 \in \mathbb{R}^{nm^2}$,

$$\begin{aligned} \langle v^2, (I_m \otimes D_{v^1} f_1) v^2 \rangle &\geq \alpha_1 \|v^2\|^2, \\ \langle v^2, ((D\gamma_1)^T \otimes I_{nm}) v^2 \rangle &\leq \ell_1 \|v^2\|^2, \end{aligned}$$

where $D_{v^1} f_1$ is evaluated at $(h_1(\sigma_2 \zeta_0^2), \sigma_2 \zeta_0^2)$ and $D\gamma_1$, α_1 , and ℓ_1 are all evaluated at $\sigma_2 \zeta_0^2$. Then (3.21) follows. \square

Lemma 3.9. *For any $\zeta_0^2 \in K_2$ and any $\zeta^2 \in \mathbb{R}^m$,*

$$(3.22) \quad \langle \zeta^2, D\gamma_2(\zeta_0^2) \zeta^2 \rangle \leq \ell_2(\zeta_0^2) \|\zeta^2\|^2.$$

Lemma 3.10. *There exists a sufficiently small σ_2 such that for any $\zeta^2 \in K_2$,*

$$(3.23) \quad \alpha_2(\zeta^2) \geq \ell_2(\zeta^2) + \sigma_2 L_2 + \frac{1}{2} c_r.$$

Proof. Since (1.7c) holds for some $r \geq 3$, we have that for any $z \in K_0$,

$$(3.24) \quad \alpha(h_0(z), z) \geq 3\ell(h_0(z), z) + 4\eta \|D_a g(h_0(z), z)\| + c_r.$$

In addition, since $L_2 > 0$ is independent of the choice of σ_2 (see Lemma 3.7), we can choose σ_2 sufficiently small so that $\sigma_2 L_2 \leq \frac{1}{2} c_r$. Combine this with (3.24). Recall (3.7), (3.9), (3.19), and (3.20), and rearrange terms accordingly. Then we obtain (3.23) once replacing z with $\sigma_1 \sigma_2 \zeta^2$. \square

Again, since the ζ^2 -component of (3.16) is independent of v^2 , Lemmas 3.7–3.10 verify that the restriction of (3.16) to $J_2 \times K_2$ satisfies Hypotheses 2.

Recall that $\|Dh_1(\zeta^1)\| < 1$ for all $\zeta^1 \in K_1$ (see Lemma 3.6). Define $h_2 : K_2 \rightarrow J_2$ as follows:

$$(3.25) \quad h_2(\zeta^2) := \text{vec}_{nm,m}(Dh_1(\sigma_2 \zeta^2)).$$

Now consider the set

$$\Gamma_2 := \{(h_2(\zeta^2), \zeta^2) : \zeta^2 \in K_2\} \subset J_2 \times K_2.$$

Clearly, Γ_2 is positively invariant under the flow of (3.16). Then, by applying Theorem 1.1 again, we obtain the C^1 smoothness of h_2 .

Lemma 3.11. Γ_2 is a C^1 manifold. In particular, $h_2 : K_2 \rightarrow J_2$ is C^1 with $\|Dh_2(\zeta^2)\| < 1$ for all $\zeta^2 \in K_2$.

Since $Dh_1(\zeta^1) = \text{vec}_{nm,m}^{-1}(h_2(\sigma_2^{-1}\zeta^1))$ for any $\zeta^1 \in K_1$, we have proven that $h_1 : K_1 \rightarrow J_1$ is C^2 with its first and second derivatives bounded on K_1 . Therefore, $h_0 : K_0 \rightarrow \mathbb{R}^n$ is now C^3 with its first to third derivatives bounded on K_0 . Furthermore,

$$D^3h_0(z)(z_1, z_2, z_3) = \sigma_2^{-1}\sigma_1^{-2} \text{vec}_{n,m}^{-1}(\text{vec}_{nm,m}^{-1}(Dh_2(\sigma_2^{-1}\sigma_1^{-1}z)z_1)z_2)z_3$$

for any $z \in K_0$ and any z_1, z_2 , and $z_3 \in \mathbb{R}^m$.

3.3.1. Induction. Up to this point, the inductive scheme of the proof has become obvious. Consider $r \geq 4$. Suppose that we have proven the C^k smoothness of h_0 for some $3 \leq k \leq r-1$ and are about to prove the C^{k+1} smoothness of h_0 . Then we have already established the following prerequisites.

- (1) For $j = 2, \dots, k$, we have defined ζ^j , K_j , and $h_j : K_j \rightarrow \mathbb{R}^{nm^j}$ recursively as follows:

$$(3.26) \quad \begin{aligned} \zeta^j &:= \sigma_j^{-1}\zeta^{j-1}, \\ K_j &:= \{\zeta^j \in \mathbb{R}^m : \sigma_j\zeta^j \in K_{j-1}\}, \\ h_j(\zeta^j) &:= \text{vec}_{nm^{j-1},m}(Dh_{j-1}(\sigma_j\zeta^j)), \end{aligned}$$

with each σ_j chosen sufficiently small. In addition, we have already proven that for each $j = 1, \dots, k-1$, h_j is C^{k-j} with $\|Dh_j(\zeta^j)\| < 1$ for all $\zeta^j \in K_j$ and all applicable higher derivatives bounded on K_j .

- (2) For $j = 2, \dots, k$, we have defined $\gamma_j : K_j \rightarrow \mathbb{R}^m$ and $f_j : \mathbb{R}^{nm^j} \times K_j \rightarrow \mathbb{R}^{nm^j}$ recursively as follows:

$$\begin{aligned} \gamma_j(\zeta^j) &:= \frac{1}{\sigma_j}\gamma_{j-1}(\sigma_j\zeta^j), \\ f_j(v^j, \zeta^j) &:= (I_m \otimes D_{v^{j-1}}f_{j-1}(h_{j-1}(\sigma_j\zeta^j), \sigma_j\zeta^j) - (D\gamma_{j-1}(\sigma_j\zeta^j))^T \otimes I_{nm^{j-1}})v^j \\ &\quad + \text{vec}_{nm^{j-1},m}(D_{\zeta^{j-1}}f_{j-1}(h_{j-1}(\sigma_j\zeta^j), \sigma_j\zeta^j)), \end{aligned}$$

where $v^j \in \mathbb{R}^{nm^j}$ and $\zeta^j \in K_j$.

- (3) For $j = 2, \dots, k$, we have defined continuous functions $\alpha_j : K_j \rightarrow \mathbb{R}$ and $\ell_j : K_j \rightarrow \mathbb{R}$ recursively as follows:

$$\begin{aligned} \alpha_j(\zeta^j) &:= \alpha_{j-1}(\sigma_j\zeta^j) - \ell_{j-1}(\sigma_j\zeta^j), \\ \ell_j(\zeta^j) &:= \ell_{j-1}(\sigma_j\zeta^j). \end{aligned}$$

Then we shall consider the system

$$(3.27) \quad \begin{aligned} \dot{v}^k &= f_k(v^k, \zeta^k), \\ \dot{\zeta}^k &= \gamma_k(\zeta^k), \end{aligned}$$

which is defined on $\mathbb{R}^{nm^k} \times K_k$. Next, choose J_k , a bounded, open subset of \mathbb{R}^{nm^k} , as follows:

$$J_k := \{v^k \in \mathbb{R}^{nm^k} : \|\text{vec}_{nm^k-1,m}^{-1}(v^k)\| < 1\}.$$

Obviously $J_k \times K_k$ satisfies Hypothesis 1. In addition, the restriction of (3.27) to $J_k \times K_k$ again satisfies Hypothesis 2, and the verification is now routine. In

particular, following the same arguments as used in the proofs of Lemmas 3.7–3.10, it is straightforward to verify the following statements about f_k , γ_k , α_k , and ℓ_k .

- (1) f_k is C^1 on $J_k \times K_k$, and γ_k is C^k on K_k . Furthermore, there exists a constant $L_k > 0$ independent of the choice of σ_k such that $\|D_{\zeta^k} f_k(v^k, \zeta^k)\| \leq \sigma_k L_k$ for any $(v^k, \zeta^k) \in J_k \times K_k$.
 - (2) For any $(v_0^k, \zeta_0^k) \in J_k \times K_k$ and any $v^k \in \mathbb{R}^{nm^k}$,
$$\langle v^k, D_{v^k} f_k(v_0^k, \zeta_0^k) v^k \rangle \geq \alpha_k(\zeta_0^k) \|v^k\|^2.$$
 - (3) For any $\zeta_0^k \in K_k$ and any $\zeta^k \in \mathbb{R}^m$,
$$\langle \zeta^k, D\gamma_k(\zeta_0^k) \zeta^k \rangle \leq \ell_k(\zeta_0^k) \|\zeta^k\|^2.$$
 - (4) There exists a sufficiently small σ_k such that for any $\zeta^k \in K_k$,
- (3.28)
$$\alpha_k(\zeta^k) \geq \ell_k(\zeta^k) + \sigma_k L_k + \frac{1}{2} c_r.$$

Now we consider the set

$$\Gamma_k := \{(h_k(\zeta^k), \zeta^k) : \zeta^k \in K_k\} \subset J_k \times K_k,$$

which, by our construction, is positively invariant under the flow of (3.27). By Theorem 1.1, Γ_k is a C^1 manifold. In particular, $h_k : K_k \rightarrow J_k$ is C^1 with $\|Dh_k(\zeta^k)\| < 1$ for all $\zeta^k \in K_k$. Then, in view of the inverse of the recurrence relation (3.26), we have proven that for each $j = 1, \dots, k$, h_j is C^{k+1-j} with $\|Dh_j(\zeta^j)\| < 1$ for all $\zeta^j \in K_j$ and all applicable higher derivatives bounded on K_j . Therefore, $h_0 : K_0 \rightarrow \mathbb{R}^n$ is now C^{k+1} with its first to $(k+1)$ -th derivatives bounded on K_0 . Furthermore,

$$\begin{aligned} D^{k+1} h_0(z)(z_1, \dots, z_{k+1}) \\ = \sigma_k^{-1} \cdots \sigma_1^{-k} \text{vec}_{n,m}^{-1}(\cdots \text{vec}_{nm^{k-1},m}^{-1}(Dh_k(\sigma_k^{-1} \cdots \sigma_1^{-1} z) z_1) z_2 \cdots) z_{k+1} \end{aligned}$$

for any $z \in K_0$ and any $z_1, \dots, z_{k+1} \in \mathbb{R}^m$.

Finally, we remark that the inequality (3.28) is guaranteed by (1.7c) provided that $k+1 \leq r$. Therefore, the induction has to cease at $k = r-1$ with the outcome of the last iteration being that h_0 is $C^{k+1} = C^r$. This concludes the proof of Theorem 1.2.

4. APPLICATIONS

4.1. Invariant Tori in a 2-Parameter Family of Systems. Consider the following 2-parameter family of systems:

$$\begin{aligned} \dot{R} &= R(8\beta - R) + \sin \phi_1 + \sin \phi_2, \\ \dot{\phi}_1 &= \beta R + \beta^2 \sin \phi_1 \sin \phi_2, \\ \dot{\phi}_2 &= \omega, \end{aligned} \tag{4.1}$$

where $R \in \mathbb{R}$, $\phi_1 \in [0, 2\pi]$ and $\phi_2 \in [0, 2\pi]$ are two angular variables with the end points 0 and 2π identified, and $\beta \in (0, \infty)$ and $\omega \in \mathbb{R}$ are the two parameters. We are interested in determining for what parameter values (4.1) has C^r ($r \geq 1$) invariant tori. However, even with the help of numerical methods, it is a challenging task to identify the exact set of parameters for the existence of C^r invariant tori. Here we apply Theorem 1.1 and Theorem 1.3 to derive sufficient conditions on the parameters for the existence of C^r invariant tori for (4.1).

We need to modify (4.1) slightly to fit the form of (1.3). This is done by lifting ϕ_1 and ϕ_2 to \mathbb{R} . Furthermore, we rescale ϕ_1 to $\theta_1 := k^{-1}\phi_1$ and ϕ_2 to $\theta_2 := (k\gamma)^{-1}\phi_2$ with the scaling factors $k > 0$ and $\gamma > 0$ to be determined later. The purpose of this rescaling of variables is to obtain a larger set of parameters for which the existence of C^r invariant tori can be guaranteed by Theorem 1.1 and Theorem 1.3. Write $\theta = (\theta_1, \theta_2)$. Then (4.1) becomes

$$(4.2) \quad \begin{aligned} \dot{R} &= f(R, \theta) := R(8\beta - R) + \sin k\theta_1 + \sin k\gamma\theta_2, \\ \dot{\theta} &= g(R, \theta) := \begin{pmatrix} \frac{\beta}{k}R + \frac{\beta^2}{k} \sin k\theta_1 \sin k\gamma\theta_2 \\ \frac{\omega}{k\gamma} \end{pmatrix}, \end{aligned}$$

where we have treated $R \in \mathbb{R}$ and $\theta \in \mathbb{R}^2$ as “ a ” and “ z ” of (1.3), respectively. In addition, for any $\delta > 0$, define the set U_δ as follows:

$$U_\delta := \{(R, \theta) : |R| < \delta, \theta \in \mathbb{R}^2\},$$

which clearly satisfies Hypothesis 1.

In order to apply Theorem 1.1, we have to first establish the existence of a positively invariant set Γ that is contained in U_δ and satisfies

$$(4.3) \quad \Pi_\perp(\Gamma) = \Pi_\perp(U_\delta) = \mathbb{R}^2,$$

where Π_\perp is the projection onto the θ -coordinate. We define $\Gamma \subset U_\delta$ to be the set of points that stay in U_δ forever along solution trajectories of (4.2) in forward time. Obviously, Γ is the largest positively invariant subset of U_δ by this definition. Next, we show that Γ satisfies (4.3) using the Ważewski theorem.

Note that at the boundary of U_δ ,

$$\begin{aligned} \dot{R} &\geq \delta(8\beta - \delta) - 2 && \text{for any } (R, \theta) \text{ with } R = \delta, \\ \dot{R} &\leq -\delta(8\beta + \delta) + 2 && \text{for any } (R, \theta) \text{ with } R = -\delta. \end{aligned}$$

Thus, for any $\beta > 0$ and $\delta > 0$ satisfying

$$(4.4) \quad \delta(8\beta - \delta) - 2 > 0,$$

$\dot{R} > 0$ for any (R, θ) with $R = \delta$, and $\dot{R} < 0$ for any (R, θ) with $R = -\delta$. Then it can be easily verified that $\text{cl}(U_\delta)$ is a Ważewski set with its boundary ∂U_δ being the set of points which leave $\text{cl}(U_\delta)$ immediately along solution trajectories of (4.2) in forward time. Let $U_\delta^0 \subseteq \text{cl}(U_\delta)$ be the set of points that do not stay in $\text{cl}(U_\delta)$ forever in forward time. Then by the Ważewski theorem, there exists a continuous function $\mathcal{R} : U_\delta^0 \times [0, 1] \rightarrow U_\delta^0$ such that \mathcal{R} is a strong deformation retraction of U_δ^0 onto ∂U_δ . Now suppose that for some θ^* ,

$$\Gamma \cap \{(R, \theta^*) : |R| \leq \delta\} = \emptyset.$$

Then the set $\{(R, \theta^*) : |R| \leq \delta\}$ is contained inside the domain of the continuous function $\mathcal{R}(\cdot, 1) : U_\delta^0 \rightarrow \partial U_\delta$. Immediately, we obtain a continuous function defined on $[-\delta, \delta]$ as follows:

$$R \mapsto \Pi \circ \mathcal{R}((R, \theta^*), 1),$$

where Π is the projection onto the R -coordinate. However, the above continuous function has the impossible property that

$$\begin{aligned} \Pi \circ \mathcal{R}((R, \theta^*), 1) &\in \{-\delta, \delta\}, \\ \Pi \circ \mathcal{R}((\delta, \theta^*), 1) &= \delta, \end{aligned}$$

$$\Pi \circ \mathcal{R}((-\delta, \theta^*), 1) = -\delta.$$

Thus, $\Gamma \cap \{(R, \theta) : |R| \leq \delta\} \neq \emptyset$ for any $\theta \in \mathbb{R}^2$. This confirms (4.3).

Next, we verify Hypothesis 2 and Hypothesis 2* for (4.2) on U_δ . Straightforward calculation shows that for any $(R, \theta) \in U_\delta$ and any $(R', \theta') \in \mathbb{R} \times \mathbb{R}^2$,

$$\begin{aligned} \langle R', D_R f(R, \theta) R' \rangle &\geq (8\beta - 2\delta) |R'|^2, \\ \langle \theta', D_\theta g(R, \theta) \theta' \rangle &\leq \beta^2 \sqrt{1 + \gamma^2} \|\theta'\|^2, \\ \|D_\theta f(R, \theta)\| &\leq k \sqrt{1 + \gamma^2}, \\ \|D_R g(R, \theta)\| &= \frac{\beta}{k}. \end{aligned}$$

It follows that for any $\beta > 0$, $\delta > 0$, and $k > 0$ satisfying

$$(4.5) \quad 8\beta - 2\delta - \beta^2 - k - \frac{\beta}{k} > 0,$$

there exist sufficiently small $\gamma > 0$ and $c_1 > 0$, both depending on β , δ , and k , such that

$$8\beta - 2\delta \geq \beta^2 \sqrt{1 + \gamma^2} + k \sqrt{1 + \gamma^2} + \frac{\beta}{k} + c_1,$$

which implies (1.6c). Therefore, by Theorem 1.1, (4.4) and (4.5) together guarantee that Γ is a C^1 positively invariant manifold. Furthermore, if there exists an integer $r \geq 2$ such that β , δ , and k also satisfy

$$(4.6) \quad 8\beta - 2\delta - r\beta^2 - (r+1)\frac{\beta}{k} > 0,$$

then, taking a sufficiently small $c_r > 0$ and reducing γ further if necessary, we have

$$8\beta - 2\delta \geq r\beta^2 \sqrt{1 + \gamma^2} + (r+1)\frac{\beta}{k} + c_r,$$

which implies (1.6d). Consequently, if β , δ , and k satisfy (4.4)–(4.6) for some $r \geq 2$, then Γ is in fact a C^r manifold by Theorem 1.3.

Note that the inequalities (4.4)–(4.6) involve the auxiliary parameters δ and k , which do not appear in the original system (4.1). In order to eliminate these auxiliary parameters to obtain conditions on β only, we consider the β -projection of the solution set of (4.4)–(4.6). Let Q_1 denote the set of triples (β, δ, k) that satisfy both (4.4) and (4.5). In addition, for $r = 2, 3, 4, \dots$, let Q_r denote the set of triples (β, δ, k) that satisfy (4.4), (4.5), and (4.6) for the corresponding r . Elementary computation reveals that $Q_r = \emptyset$ for all $r \geq 8$. Furthermore, we find that for each $r = 1, \dots, 7$, the β -projection of Q_r is a connected interval (β_r^-, β_r^+) with the approximate values of β_r^- and β_r^+ given in the table below. Note that $(\beta_1^-, \beta_1^+) \supset (\beta_2^-, \beta_2^+) \supset \dots \supset (\beta_7^-, \beta_7^+)$.

r	1	2	3	4	5	6	7
β_r^-	0.395	0.404	0.420	0.441	0.472	0.518	0.634
β_r^+	7.248	3.887	2.542	1.849	1.412	1.093	0.781

TABLE 1. β_r^- and β_r^+ .

Now consider (4.1) with an arbitrary $\omega \in \mathbb{R}$ and a β drawn from one of the intervals (β_r^-, β_r^+) , $r = 1, \dots, 7$. Since (β_r^-, β_r^+) is the β -projection of Q_r , we can always find appropriate δ , k , and correspondingly a sufficiently small γ such that Theorem 1.1 (if $r = 1$) or Theorem 1.3 (if $2 \leq r \leq 7$) is applicable to the modified system (4.2) with $(\beta, \omega, k, \gamma)$ on the domain U_δ . Then we establish the existence of the C^r positively invariant manifold Γ , which, however, depends on β and ω as

well as the choice of the auxiliary parameters k , γ , and δ . Specifically, by Theorem 1.1 or Theorem 1.3, Γ is the graph of a C^r function $h(\cdot, \cdot; \beta, \omega, k, \gamma) : \mathbb{R}^2 \rightarrow (-\delta, \delta)$, i.e.,

$$\Gamma = \{(h(\theta_1, \theta_2; \beta, \omega, k, \gamma), \theta_1, \theta_2) : (\theta_1, \theta_2) \in \mathbb{R}^2\} \subset U_\delta.$$

Consequently, for (4.1) with (ϕ_1, ϕ_2) lifted to \mathbb{R}^2 , we obtain a C^r positively invariant manifold

$$(4.7) \quad \{(h(\frac{\phi_1}{k}, \frac{\phi_2}{k\gamma}; \beta, \omega, k, \gamma), \phi_1, \phi_2) : (\phi_1, \phi_2) \in \mathbb{R}^2\}.$$

On the other hand, suppose we can choose k' , γ' , and δ' such that for (4.2) with $(\beta, \omega, k', \gamma')$ there is also a C^r positively invariant manifold Γ' , which is the graph of a C^r function $h(\cdot, \cdot; \beta, \omega, k', \gamma') : \mathbb{R}^2 \rightarrow (-\delta', \delta')$, i.e.,

$$\Gamma' = \{(h(\theta'_1, \theta'_2; \beta, \omega, k', \gamma'), \theta'_1, \theta'_2) : (\theta'_1, \theta'_2) \in \mathbb{R}^2\} \subset U_{\delta'}.$$

Then, for (4.1) with (ϕ_1, ϕ_2) lifted to \mathbb{R}^2 , we obtain “another” C^r positively invariant manifold

$$\{(h(\frac{\phi_1}{k'}, \frac{\phi_2}{k'\gamma'}; \beta, \omega, k', \gamma'), \phi_1, \phi_2) : (\phi_1, \phi_2) \in \mathbb{R}^2\}.$$

We show that in fact

$$(4.8) \quad h(\frac{\phi_1}{k}, \frac{\phi_2}{k\gamma}; \beta, \omega, k, \gamma) = h(\frac{\phi_1}{k'}, \frac{\phi_2}{k'\gamma'}; \beta, \omega, k', \gamma') \text{ for all } (\phi_1, \phi_2) \in \mathbb{R}^2.$$

Without loss of generality, we assume $\delta \leq \delta'$. It follows that for all $(\theta_1, \theta_2) \in \mathbb{R}^2$,

$$|h(\theta_1, \theta_2; \beta, \omega, k, \gamma)| < \delta'.$$

In addition, by the relation $(\phi_1, \phi_2) = (k\theta_1, k\gamma\theta_2) = (k'\theta'_1, k'\gamma'\theta'_2)$, the positive invariance of Γ under the flow of (4.2) with $(\beta, \omega, k, \gamma)$ implies that

$$\tilde{\Gamma} := \{(h(\frac{k'\theta'_1}{k}, \frac{k'\gamma'\theta'_2}{k\gamma}; \beta, \omega, k, \gamma), \theta'_1, \theta'_2) : (\theta'_1, \theta'_2) \in \mathbb{R}^2\} \subset U_{\delta'}$$

is positively invariant under the flow of (4.2) with $(\beta, \omega, k', \gamma')$. Recall that by definition Γ' is the largest positively invariant subset of $U_{\delta'}$ for (4.2) with $(\beta, \omega, k', \gamma')$. Thus, $\tilde{\Gamma} \subseteq \Gamma'$, which implies that for all $(\theta'_1, \theta'_2) \in \mathbb{R}^2$,

$$h(\frac{k'\theta'_1}{k}, \frac{k'\gamma'\theta'_2}{k\gamma}; \beta, \omega, k, \gamma) = h(\theta'_1, \theta'_2; \beta, \omega, k', \gamma').$$

Then we obtain (4.8) by substituting $\theta'_1 = \frac{\phi_1}{k'}$ and $\theta'_2 = \frac{\phi_2}{k'\gamma'}$ into the above identity. Therefore, for each $(\beta, \omega) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}$, (4.1) with (ϕ_1, ϕ_2) lifted to \mathbb{R}^2 has a unique C^r positively invariant manifold that is contained in U_δ for a certain $\delta > 0$ and that is the graph of a C^r function from \mathbb{R}^2 to \mathbb{R} as given by (4.7).

We still need to show that (4.7) corresponds to a C^r invariant torus. By the definition of Γ , the trajectory that passes through $(h(\theta_1, \theta_2; \beta, \omega, k, \gamma), \theta_1, \theta_2)$ is contained in U_δ forever in forward time. Then by the periodicity of (4.2) in θ_1 and θ_2 , the trajectory that passes through $(h(\theta_1, \theta_2; \beta, \omega, k, \gamma), \theta_1 + \frac{2n\pi}{k}, \theta_2 + \frac{2m\pi}{k\gamma})$ for any integers n and m is also contained in U_δ forever in forward time. Thus $(h(\theta_1, \theta_2; \beta, \omega, k, \gamma), \theta_1 + \frac{2n\pi}{k}, \theta_2 + \frac{2m\pi}{k\gamma}) \in \Gamma$, which implies that

$$(4.9) \quad h(\theta_1, \theta_2; \beta, \omega, k, \gamma) = h(\theta_1 + \frac{2n\pi}{k}, \theta_2 + \frac{2m\pi}{k\gamma}; \beta, \omega, k, \gamma).$$

For $(\phi_1, \phi_2) \in \mathbb{R}^2$, let $\rho^u(\phi_1, \phi_2; \beta, \omega) := h(\frac{\phi_1}{k}, \frac{\phi_2}{k\gamma}; \beta, \omega, k, \gamma)$, which is well defined in view of (4.8). Clearly, $\rho^u(\phi_1, \phi_2; \beta, \omega)$ is C^r in ϕ_1 and ϕ_2 , and it is 2π -periodic in ϕ_1 and ϕ_2 by (4.9). Therefore, (4.7) corresponds to the C^r torus

$$T_{\beta, \omega}^u := \{(\rho^u(\phi_1, \phi_2; \beta, \omega), \phi_1, \phi_2) : \phi_1, \phi_2 \in \mathbb{R}(\text{mod } 2\pi)\}.$$

It is easy to verify that $T_{\beta,\omega}^u$ is invariant under the flow of (4.1) (with the corresponding β and ω) in both forward time and backward time. Note that the superscript “ u ” reflects the obvious fact that $T_{\beta,\omega}^u$ is unstable.

After establishing $T_{\beta,\omega}^u$, we also consider the time reversal of (4.2) on the domain $\{(R, \theta) : |R - 8\beta| < \delta, \theta \in \mathbb{R}^2\}$. By repeating all the analysis presented so far in this subsection, we find that for any $(\beta, \omega) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}$ with $1 \leq r \leq 7$, (4.1) has another C^r invariant torus

$$T_{\beta,\omega}^s := \{(\rho^s(\phi_1, \phi_2; \beta, \omega), \phi_1, \phi_2) : \phi_1, \phi_2 \in \mathbb{R}(\text{mod } 2\pi)\},$$

where $\rho^s(\phi_1, \phi_2; \beta, \omega)$ is C^r and 2π -periodic in ϕ_1 and ϕ_2 for $(\phi_1, \phi_2) \in \mathbb{R}^2$. $T_{\beta,\omega}^s$ is stable and contained in the domain $\{(R, \theta) : |R - 8\beta| < \delta, \theta \in \mathbb{R}^2\}$ for a certain $\delta > 0$.

Furthermore, for any $(\beta_0, \omega_0) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}$ with $1 \leq r \leq 7$, let $\beta = \beta_0 + \epsilon\xi_1$ and $\omega = \omega_0 + \epsilon\xi_2$ for a sufficiently small $\epsilon > 0$ such that $\beta = \beta_0 + \epsilon\xi_1 \in (\beta_r^-, \beta_r^+)$ for all $|\xi_1| < 1$. Next, we append $\dot{\xi}_1 = 0$ and $\dot{\xi}_2 = 0$ to (4.1) to form an enlarged system with the additional components ξ_1 and ξ_2 . Then by adapting our previous analysis for the enlarged system, reducing ϵ further if necessary, we obtain that the enlarged system has C^r invariant manifolds M^u and M^s which are the graphs of the C^r functions, respectively,

$$\begin{aligned} (\phi_1, \phi_2, \xi_1, \xi_2) &\mapsto \rho^u(\phi_1, \phi_2; \beta_0 + \epsilon\xi_1, \omega_0 + \epsilon\xi_2) \text{ and} \\ (\phi_1, \phi_2, \xi_1, \xi_2) &\mapsto \rho^s(\phi_1, \phi_2; \beta_0 + \epsilon\xi_1, \omega_0 + \epsilon\xi_2) \end{aligned}$$

on the domain $\mathbb{R}(\text{mod } 2\pi) \times \mathbb{R}(\text{mod } 2\pi) \times (-1, 1) \times (-1, 1)$. It follows that $\rho^u(\phi_1, \phi_2; \beta, \omega)$ and $\rho^s(\phi_1, \phi_2; \beta, \omega)$ are also C^r in β and ω for $(\beta, \omega) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}$. Therefore, for each $r = 1, \dots, 7$, $\{T_{\beta,\omega}^u : (\beta, \omega) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}\}$ and $\{T_{\beta,\omega}^s : (\beta, \omega) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}\}$ are C^r families of C^r invariant tori.

Finally, we compare our estimates (i.e., β_r^- and β_r^+ in Table 1) to some preliminary numerical results:

- (1) By numerically integrating (4.1), we identify a region on the (β, ω) -plane bounded between the line $\beta = 0$ and the curve S_0 (see Figure 1) such that for (4.1) with any (β, ω) in this region, solution trajectories that start on the plane $\{(R, \phi_1, \phi_2) : R = 5\}$ all reach the plane $\{(R, \phi_1, \phi_2) : R = -5\}$. Note that for (4.1) with $\beta < \beta_1^-$, $\dot{R} < 0$ if $R \leq -5$ or $R \geq 5$. Thus, we conclude that no invariant torus exists for (4.1) with any (β, ω) in this region. In addition, when we continue the tori $T_{\beta,\omega}^u$ and $T_{\beta,\omega}^s$ along a path with constant ω starting at a β between β_1^- and β_2^- and moving towards the curve S_0 , we observe that $T_{\beta,\omega}^u$ and $T_{\beta,\omega}^s$ continue to exist even for $\beta < \beta_1^-$, but as (β, ω) gets close to S_0 , both tori go through a series of bifurcations and then disappear. We have not yet studied these bifurcations in detail since it requires a much more delicate numerical analysis, which is out of the scope of this work.
- (2) By sweeping the region $\beta_7^+ \leq \beta \leq 9$ and $-2 \leq \omega \leq 2$, we find that for some (β, ω) , the α -limit set of $T_{\beta,\omega}^u$ consists of a saddle-type periodic orbit P_1 and a non-saddle-type unstable periodic orbit P_2 . In addition, the tangent bundle of \mathbb{R}^3 restricted to P_2 has the splitting: $T\mathbb{R}^3|_{P_2} = N_1 \oplus N_2 \oplus TP_2$ such that both N_1 and N_2 are invariant under the linearized flow of (4.1) along P_2 and $T_{\beta,\omega}^u$ is tangent to N_2 along P_2 . Since the linearized flow

along P_2 expands both N_1 and N_2 , the ratio between the expansion rate in N_1 and the expansion rate in N_2 determines the smoothness of $T_{\beta,\omega}^u$. By continuing $T_{\beta,\omega}^u$ and P_2 and monitoring the expansion rates in N_1 and N_2 , we obtain the level curves S_2, \dots, S_9 (see Figure 2), along which the ratio between the expansion rates in N_1 and N_2 are integers 2, ..., 9, respectively. In addition, we find that for any (β, ω) on the curve S_1 , the bundles N_1 and N_2 merge, and then for (β, ω) on the right-hand side of S_1 , the two bundles emerge again, however, both rotating $n\pi$ ($n \in \mathbb{Z} \setminus \{0\}$) for one period of P_2 . Thus, $T_{\beta,\omega}^u$ does not exist for any (β, ω) on the right-hand side of S_1 , and $T_{\beta,\omega}^u$ is exactly C^r (i.e., not C^{r+1}) for any (β, ω) in the region bounded between S_{r+1} and S_r ($r = 1, \dots, 8$). We remark that $T_{\beta,\omega}^u$ can be C^r with $r \geq 10$ for some (β, ω) on the left-hand side of S_9 .

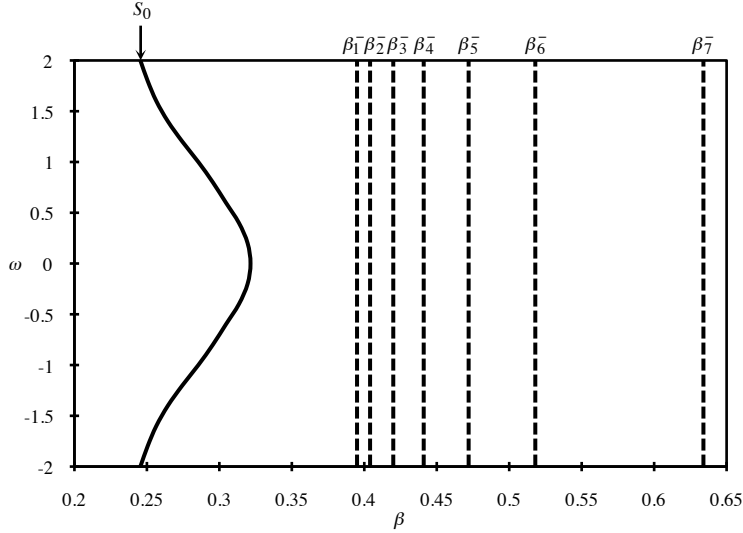


FIGURE 1. β_r^- and the curve S_0 on the (β, ω) -plane. No invariant torus exists for (4.1) with any (β, ω) in the region bounded between the line $\beta = 0$ and the curve S_0 .

4.2. Invariant Torus in a System with Rapid Oscillations. In [5], Chicone and Liu studied the existence of an invariant torus in the following system:

$$\begin{aligned}
 \dot{\rho} &= \Delta(\sigma)\rho + \Lambda(\sigma) + \mu R(\rho, \sigma, \tau/\mu^2, \mu), \\
 \dot{\sigma} &= \rho + \mu S(\rho, \sigma, \tau/\mu^2, \mu), \\
 \dot{\tau} &= 1,
 \end{aligned}
 \tag{4.10}$$

where $\dot{\cdot} = \frac{d}{ds}$ for a slow time s , $\mu > 0$ is a real parameter, $\rho \in \mathbb{R}$, and $\sigma \in [0, 2\pi]$ and $\tau \in [0, 2\pi\mu^2/\omega]$ for a fixed $\omega > 0$ are angular variables with the corresponding end points identified. In addition, (4.10) is assumed to be class C^∞ . Note that μR and μS are fast oscillatory in τ for μ small.

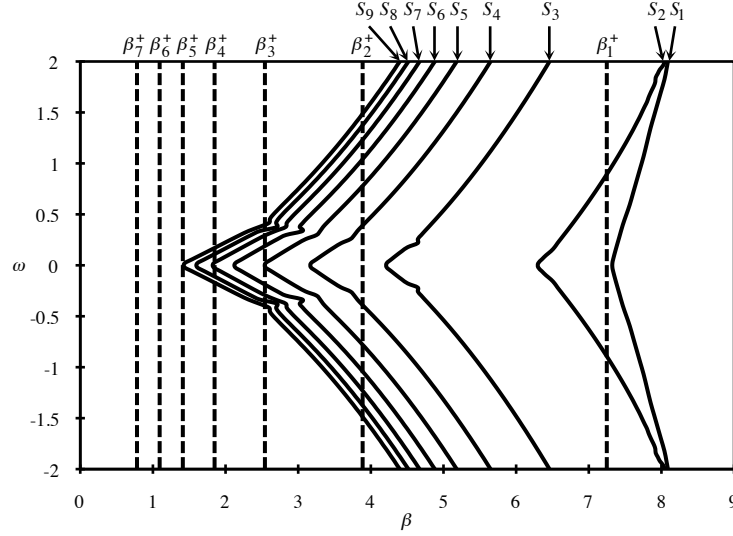


FIGURE 2. β_r^+ and the curves S_i on the (β, ω) -plane. $T_{\beta, \omega}^u$ does not exist for any (β, ω) on the right-hand side of S_1 , and $T_{\beta, \omega}^u$ is exactly C^r (i.e., not C^{r+1}) for any (β, ω) in the region bounded between S_{r+1} and S_r ($r = 1, \dots, 8$). Note that our theoretical estimates predict that $T_{\beta, \omega}^u$ is (at least) C^r for any $(\beta, \omega) \in (\beta_r^-, \beta_r^+) \times \mathbb{R}$ with $1 \leq r \leq 7$.

After the truncation of the $\mathcal{O}(\mu)$ terms μR and μS , (4.10) becomes

$$(4.11) \quad \begin{aligned} \dot{\rho} &= \Delta(\sigma)\rho + \Lambda(\sigma), \\ \dot{\sigma} &= \rho, \\ \dot{\tau} &= 1. \end{aligned}$$

It is assumed in [5] that $\Lambda(\sigma) \neq 0$ and $\Delta(\sigma) < 0$ for all $\sigma \in [0, 2\pi]$. Then the ρ - σ subsystem of (4.11) has an attracting limit cycle, whose existence can be proved by an easy application of the Poincaré-Bendixson theorem, and the suspension of this limit cycle in (4.11) forms an invariant torus \mathcal{M}_0 . An immediate question is whether or not \mathcal{M}_0 persists in the system (4.10) for $\mu > 0$ but small. Note that the attracting torus \mathcal{M}_0 is r -normally hyperbolic with respect to (4.11) for any integer $r \geq 1$ and its strength of normal hyperbolicity is independent of any parameter. However, the persistence theory of r -normally hyperbolic invariant manifolds by Fenichel [9] and Hirsch, Pugh, and Shub [14] is not applicable in this case. The reason is that (4.10) may not be C^1 close to (4.11) even for small μ since the partial derivatives of μR and μS with respect to t can be very large for small μ . On the other hand, by rescaling time to $t := s/\mu^2$ and taking $\theta := \tau/\mu^2$, we transform (4.10) into (1.13) with $w = (\rho, \sigma)$ and $\epsilon = \mu^2$. Then, in the $\mathcal{O}(\epsilon)$ truncation of (1.13) (i.e., $\dot{w} = \epsilon F_1(w)$, $\dot{\theta} = 1$), the normal hyperbolicity of the corresponding unperturbed invariant manifold depends on the small parameter ϵ , which raises the problem of weak hyperbolicity as discussed in Subsection 1.2.

The main result of Chicone and Liu is that *for any integer $r \geq 2$, (4.10) with $\mu > 0$ but sufficiently small has an r -normally hyperbolic invariant torus C^r diffeomorphic to \mathcal{M}_0* . In addition, by time reversal, the same conclusion holds under the assumption that $\Lambda(\sigma) \neq 0$ and $\Delta(\sigma) > 0$ for all $\sigma \in [0, 2\pi]$. Chicone and Liu also remarked that this result is not valid if Λ is allowed to have zeros and the formulation of correct hypotheses needed to prove an analogous result in this case is “an interesting open problem”. One of the difficulties is that even the existence of a C^r invariant torus for (4.11) is not readily available if Λ has zeros.

We now apply Theorem 1.1 and Theorem 1.3 to establish a C^r ($r \geq 1$) invariant torus for (4.10) even if Λ has zeros. In preparation for our analysis, we lift σ and τ to \mathbb{R} and then rescale them to $\theta_1 := k^{-1}\sigma$ and $\theta_2 := \tau/\mu^2$ with the scaling factor $k > 0$ to be determined later. In addition, we introduce an auxiliary variable $\theta_3 \in (-2, 2)$ to (4.10) to form an enlarged system, which is written in the form of (1.3) as follows:

$$(4.12) \quad \begin{aligned} \dot{\rho} &= f(\rho, \theta) := \Delta(k\theta_1)\rho + \Lambda(k\theta_1) + \theta_3\mu R(\rho, k\theta_1, \theta_2, \mu), \\ \dot{\theta} = g(\rho, \theta) &:= \begin{pmatrix} \frac{1}{k}\rho + \frac{1}{k}\theta_3\mu S(\rho, k\theta_1, \theta_2, \mu) \\ \frac{1}{\mu^2} \\ 0 \end{pmatrix}. \end{aligned}$$

where $\theta := (\theta_1, \theta_2, \theta_3)$, Δ and Λ are $2\pi/k$ -periodic in θ_1 , and R and S are $2\pi/k$ -periodic in θ_1 and $2\pi/\omega$ -periodic in θ_2 . Note that we have treated ρ and θ as “ a ” and “ z ” of (1.3), respectively.

Assume $\Delta(\sigma) > 0$ for all $\sigma \in [0, 2\pi]$. Define constants E_1 , E_2 , and E as follows:

$$(4.13) \quad \begin{aligned} E_1 &:= \min_{\sigma \in [0, 2\pi]} \left(-\frac{\Lambda(\sigma)}{\Delta(\sigma)} \right), \\ E_2 &:= \max_{\sigma \in [0, 2\pi]} \left(-\frac{\Lambda(\sigma)}{\Delta(\sigma)} \right), \\ E &:= \max \{ |E_1|, |E_2| \}. \end{aligned}$$

We shall consider (4.12) on the domain

$$U_\delta := (E_1 - \delta, E_2 + \delta) \times \mathbb{R}^2 \times (-2, 2),$$

where $\delta \in (0, 1]$ is another fixed constant to be determined. Obviously, U_δ as defined above satisfies Hypothesis 1. Now, we need to establish the existence of a positively invariant set Γ that is contained in U_δ and satisfies

$$(4.14) \quad \Pi_\perp(\Gamma) = \Pi_\perp(U_\delta) = \mathbb{R}^2 \times (-2, 2),$$

where Π_\perp is the projection onto the θ -coordinate. In the same way as we have done for the example in Subsection 4.1, we define $\Gamma \subset U_\delta$ to be the set of points that stay in U_δ forever along solution trajectories of (4.12) in forward time so that Γ is the largest positively invariant subset of U_δ by definition. Then we show that Γ satisfies (4.14), again, using the Ważewski theorem.

By the definitions of E_1 and E_2 and the assumption that $\Delta(\sigma) > 0$ for all $\sigma \in [0, 2\pi]$, it is straightforward to verify that for $M := \min_{\sigma \in [0, 2\pi]} \Delta(\sigma) > 0$,

$$\begin{aligned} \dot{\rho} &\geq M\delta - 2\mu\|R\|_0 && \text{for any } (\rho, \theta) \text{ with } \rho \geq E_2 + \delta, \\ \dot{\rho} &\leq -M\delta + 2\mu\|R\|_0 && \text{for any } (\rho, \theta) \text{ with } \rho \leq E_1 - \delta, \end{aligned}$$

where $\|R\|_0$ is the uniform norm of $R(\rho, k\theta_1, \theta_2, \mu)$ on the domain $(E_1 - 1, E_2 + 1) \times \mathbb{R}^2 \times (0, 1]$. Then for any $\delta, \mu \in (0, 1]$ satisfying

$$(4.15) \quad M\delta - 2\mu\|R\|_0 > 0,$$

$\dot{\rho} > 0$ for all (ρ, θ) with $\rho \geq E_2 + \delta$, and $\dot{\rho} < 0$ for all (ρ, θ) with $\rho \leq E_1 - \delta$. It follows that $\{E_1 - \delta, E_2 + \delta\} \times \mathbb{R}^2 \times (-2, 2)$ is the set of points which leave the set $[E_1 - \delta, E_2 + \delta] \times \mathbb{R}^2 \times (-2, 2)$ immediately along solution trajectories of (4.12) in forward time. Then, the latter is a Ważewski set, and by the same arguments as what we have used in Subsection 4.1, we can show that Γ does satisfy (4.14).

Next, we verify Hypothesis 2 and Hypothesis 2* for (4.12) on U_δ . Straightforward calculation shows that for any $(\rho, \theta) \in U_\delta$ and any $(\rho', \theta') \in \mathbb{R} \times \mathbb{R}^3$,

$$\begin{aligned} \langle \rho', D_\rho f(\rho, \theta) \rho' \rangle &\geq (\Delta(k\theta_1) - 2\mu\|D_1 R\|_0) |\rho'|^2, \\ \langle \theta', D_\theta g(\rho, \theta) \theta' \rangle &\leq \mu \left(\frac{1}{k^2} \|S\|_0^2 + 4\|D_2 S\|_0^2 + \frac{4}{k^2} \|D_3 S\|_0^2 \right)^{\frac{1}{2}} \|\theta'\|^2, \\ \|D_\theta f(\rho, \theta)\| &\leq \left(k^2 (\|D\Delta(k\theta_1)\|(E + \delta) + \|D\Lambda(k\theta_1)\| + 2\mu\|D_2 R\|_0)^2 \right. \\ &\quad \left. + 4\mu^2 \|D_3 R\|_0^2 + \mu^2 \|R\|_0^2 \right)^{\frac{1}{2}}, \\ \|D_\rho g(\rho, \theta)\| &\leq \frac{1}{k} (1 + 2\mu\|D_1 S\|_0), \end{aligned}$$

where $D_i F = \frac{\partial}{\partial x_i} F(x_1, x_2, x_3, x_4)$ for $i = 1, 2, 3$, and $F = R$, or S , and $\|\cdot\|_0$ denotes the uniform norm of functions of (x_1, x_2, x_3, x_4) on the domain $(E_1 - 1, E_2 + 1) \times \mathbb{R}^2 \times (0, 1]$. Thus, for any triple $(\mu, k, \delta) \in (0, 1] \times (0, \infty) \times (0, 1]$ such that

$$\begin{aligned} (4.16) \quad &(\Delta(k\theta_1) - 2\mu\|D_1 R\|_0) - \mu \left(\frac{1}{k^2} \|S\|_0^2 + 4\|D_2 S\|_0^2 + \frac{4}{k^2} \|D_3 S\|_0^2 \right)^{\frac{1}{2}} \\ &- \left(k^2 (\|D\Delta(k\theta_1)\|(E + \delta) + \|D\Lambda(k\theta_1)\| + 2\mu\|D_2 R\|_0)^2 + 4\mu^2 \|D_3 R\|_0^2 + \mu^2 \|R\|_0^2 \right)^{\frac{1}{2}} \\ &- \frac{1}{k} (1 + 2\mu\|D_1 S\|_0) > 0 \end{aligned}$$

holds for all $k\theta_1 \in [0, 2\pi]$, (1.6c) (with a sufficiently small c_1) and hence Hypothesis 2 are satisfied on the domain U_δ . Therefore, by Theorem 1.1, (4.15) and (4.16) together guarantee that Γ is a C^1 positively invariant manifold contained inside U_δ . Furthermore, if there also exists an integer $r \geq 2$ such that

$$(4.17) \quad \begin{aligned} &(\Delta(k\theta_1) - 2\mu\|D_1 R\|_0) - r\mu \left(\frac{1}{k^2} \|S\|_0^2 + 4\|D_2 S\|_0^2 + \frac{4}{k^2} \|D_3 S\|_0^2 \right)^{\frac{1}{2}} \\ &- (r + 1) \frac{1}{k} (1 + 2\mu\|D_1 S\|_0) > 0 \end{aligned}$$

holds for all $k\theta_1 \in [0, 2\pi]$, then (1.6d) (with a sufficiently small c_r) and hence Hypothesis 2* are satisfied on the domain U_δ . Consequently, if the triple (μ, k, δ) satisfies (4.15)–(4.17) for some $r \geq 2$, then Γ is in fact a C^r manifold by Theorem 1.3.

Note that with (4.15)–(4.17), we can determine for each positive integer r a subset of $(0, 1]$ such that the system (4.12) with any μ chosen from this subset (if nonempty) possesses a C^r positively invariant manifold Γ , which is contained inside U_δ . In principle, these subsets can be identified in the same way as we obtain Table 1 in Subsection 4.1. However, the analysis would be difficult without explicit expressions of Δ , Λ , R , and S . On the other hand, if our only concern is what happens given that μ is sufficiently small, we can derive a relatively simple

condition on Δ and Λ to guarantee the existence of the C^r ($r \geq 1$) positively invariant manifold Γ for (4.12).

First, we make some important observations: if for some $k > 0$,

$$(4.18) \quad \Delta(\sigma) > k(\|D\Delta(\sigma)\|E + \|D\Lambda(\sigma)\|) + \frac{1}{k} \text{ for all } \sigma \in [0, 2\pi],$$

then we can always find an appropriate $\delta \in (0, 1]$ and correspondingly a sufficiently small $\mu_0 > 0$ such that for any $\mu \in (0, \mu_0]$ the triple (μ, k, δ) satisfies both (4.15) and (4.16); and if in addition to (4.18) there is an integer $r \geq 2$ such that

$$\Delta(\sigma) > (r+1)\frac{1}{k} \text{ for all } \sigma \in [0, 2\pi],$$

then we can further reduce μ_0 if necessary so that for any $\mu \in (0, \mu_0]$ the triple (μ, k, δ) satisfies all (4.15)–(4.17). Next, we notice the following facts:

$$\inf_{k>0} \left\{ \max_{\sigma \in [0, 2\pi]} \left\{ k(\|D\Delta(\sigma)\|E + \|D\Lambda(\sigma)\|) + \frac{1}{k} \right\} \right\} = 2\sqrt{L},$$

$$\inf_{k>0} \left\{ \max_{\sigma \in [0, 2\pi]} \left\{ k(\|D\Delta(\sigma)\|E + \|D\Lambda(\sigma)\|) + \frac{1}{k}, (r+1)\frac{1}{k} \right\} \right\} = \frac{r+1}{\sqrt{r}}\sqrt{L},$$

where $L := \max_{\sigma \in [0, 2\pi]} \{\|D\Delta(\sigma)\|E + \|D\Lambda(\sigma)\|\}$. It follows that if

$$(4.19) \quad \Delta(\sigma) > \frac{r+1}{\sqrt{r}}\sqrt{L} \text{ for all } \sigma \in [0, 2\pi],$$

then there always exist a $k > 0$, an appropriate $\delta \in (0, 1]$, and correspondingly a sufficiently small $\mu_0 > 0$ such that for any $\mu \in (0, \mu_0]$ the triple (μ, k, δ) satisfies both (4.15) and (4.16) when $r = 1$ in (4.19) and satisfies all (4.15)–(4.17) when $r \geq 2$ in (4.19). Consequently, (4.12) with the corresponding μ and k possesses a C^r positively invariant manifold

$$\Gamma = \{(h(\theta_1, \theta_2, \theta_3; \mu, k), \theta_1, \theta_2, \theta_3) : (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^2 \times (-2, 2)\},$$

which is contained in U_δ with the corresponding δ . The function $h(\cdot, \cdot, \cdot; \mu, k) : \mathbb{R}^2 \times (-2, 2) \rightarrow (E_1 - \delta, E_2 + \delta)$ is C^r .

Although Γ depends on the choice of (μ, k, δ) for (4.12) and U_δ , we can show by the same arguments as given in Subsection 4.1 that, up to a rescaling in θ_1 , Γ coincides with any Γ' that is associated with any other admissible (μ, k', δ') as long as μ remains the same. In particular, similar to (4.8), we have that

$$h(\frac{\sigma}{k}, \theta_2, \theta_3; \mu, k) = h(\frac{\sigma}{k'}, \theta_2, \theta_3; \mu, k') \text{ for all } (\sigma, \tau) \in \mathbb{R}^2.$$

Furthermore, by the periodicity of (4.12) with respect to θ_1 and θ_2 , we have that, similar to (4.9), $h(\theta_1, \theta_2, \theta_3; \mu, k)$ is $\frac{2\pi}{k}$ -periodic in θ_1 and $\frac{2\pi}{\omega}$ -periodic in θ_2 . Then, for any $\beta \in [0, 1]$, the C^r submanifold $\Gamma|_{\theta_3=\beta}$ (i.e., the section of Γ at $\theta_3 = \beta$) corresponds to a C^r torus

$$T_{\mu, \beta} := \{(\rho(\sigma, \tau; \beta, \mu), \sigma, \tau) : \sigma \in \mathbb{R}(\bmod 2\pi), \tau \in \mathbb{R}(\bmod \frac{2\pi\mu^2}{\omega})\},$$

where $\rho(\sigma, \tau; \beta, \mu) := h(\frac{\sigma}{k}, \frac{\tau}{\mu^2}, \beta; \mu, k)$. Recall that $\dot{\rho} > 0$ for all (ρ, θ) with $\rho \geq E_2 + \delta$, $\dot{\rho} < 0$ for all (ρ, θ) with $\rho \leq E_1 - \delta$, and Γ is the largest positively invariant subset of U_δ . Thus, $T_{\mu, \beta}$ is the unique invariant torus for the following system with the corresponding μ and β :

$$\begin{aligned} \dot{\rho} &= \Delta(\sigma)\rho + \Lambda(\sigma) + \beta\mu R(\rho, \sigma, \tau/\mu^2, \mu), \\ \dot{\sigma} &= \rho + \beta\mu S(\rho, \sigma, \tau/\mu^2, \mu), \\ \dot{\tau} &= 1. \end{aligned}$$

Finally, since h is C^r with respect to θ_3 , we have that for every fixed $\mu \in (0, \mu_0]$, $\{T_{\mu, \beta} : \beta \in [0, 1]\}$ forms a C^r family of tori. Thus we have established simultaneously the existence of the invariant torus $T_{\mu, 1}$ for (4.10), the existence of the invariant torus $\mathcal{M}_0 := T_{\mu, 0}$ for (4.11), and the fact that $T_{\mu, 1}$ is C^r diffeomorphic to \mathcal{M}_0 . Notice that the existence of \mathcal{M}_0 for (4.11) is not among our assumptions. We summarize these results in the following proposition as an answer to the open question posed by Chicone and Liu in [5].

Proposition 4.1. *Suppose (4.10) is class C^∞ and $\Delta(\sigma) > 0$ for all $\sigma \in [0, 2\pi]$. Let E be defined as (4.13), and define $L := \max_{\sigma \in [0, 2\pi]} \{\|D\Delta(\sigma)\|E + \|D\Lambda(\sigma)\|\}$. If there is an integer $r \geq 1$ such that $\Delta(\sigma) > \frac{r+1}{\sqrt{r}}\sqrt{L}$ for all $\sigma \in [0, 2\pi]$, then*

- (1) (4.11) has a unique C^r invariant torus \mathcal{M}_0 ; and
- (2) there exists a sufficiently small $\mu_0 > 0$ such that for any $\mu \in (0, \mu_0]$, (4.10) has a unique C^r invariant torus $T_{\mu, 1}$, which is C^r diffeomorphic to \mathcal{M}_0 .

4.3. Persistence of a Weakly Normally Hyperbolic Invariant Torus. In this subsection, we consider the persistence of a weakly normally hyperbolic invariant torus in the following system:

$$(4.20) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w) + \epsilon^{1+\mu} F_2(w, \theta, \epsilon), \\ \dot{\theta} &= \Theta_0 + \epsilon^\nu G_1(w, \epsilon) + \epsilon^{1+\gamma} G_2(w, \theta, \epsilon), \end{aligned}$$

where $w \in \mathbb{R}^n$, $\theta \in \mathbb{T}^m$, $0 < \epsilon \ll 1$, and the power indices μ , ν , and γ satisfy

$$(4.21) \quad \begin{aligned} \mu &> 0, \\ \gamma &> 0, \\ 1 &\geq \nu \geq 0, \\ \mu + \nu &> 1. \end{aligned}$$

Assumption 4.2. F_1 is C^{r+1} ($r \geq 1$). F_2 , G_1 , and G_2 are C^r with respect to (w, θ) and continuous with respect to ϵ for $\epsilon \in [0, \epsilon_0]$.

Suppose the averaged equation $\dot{w} = \epsilon F_1(w)$ has a periodic orbit S . Note that the existence and the geometry of this periodic orbit is completely determined by the vector field $F_1(w)$ and independent of ϵ . Furthermore, the C^{r+1} smoothness of F_1 implies that S is C^{r+2} . Thus $T_0 := S \times \mathbb{T}^m$ is a C^{r+2} invariant $(m+1)$ -torus for the following truncated system with any $\epsilon > 0$:

$$(4.22) \quad \begin{aligned} \dot{w} &= \epsilon F_1(w), \\ \dot{\theta} &= \Theta_0 + \epsilon^\nu G_1(w, \epsilon), \end{aligned}$$

where $\epsilon^{1+\mu} F_2(w, \theta, \epsilon)$ and $\epsilon^{1+\gamma} G_2(w, \theta, \epsilon)$, the terms that are higher order in ϵ , are excluded. We assume that T_0 is normally hyperbolic with respect to the flow of (4.22) for any $\epsilon > 0$. This is true if and only if the periodic orbit S is hyperbolic with respect to the flow of the averaged equation $\dot{w} = \epsilon F_1(w)$ for any $\epsilon > 0$. Thus we formulate the hyperbolicity assumption as follows:

Assumption 4.3. Let $\chi(\zeta)$ with $\chi(0) \in S$ be a periodic solution of $\frac{dw}{d\zeta} = F_1(w)$, and let ζ_0 be the period of $\chi(\zeta)$. The linear variational equation $\frac{dw}{d\zeta} = DF_1(\chi(\zeta))w$ has n linearly independent solutions whose Lyapunov exponents r_1, \dots, r_n satisfy

$$\begin{aligned} \sigma_u &:= \min \{r_i : 1 \leq i \leq n_u\} > 0, \\ -\sigma_s &:= \max \{r_i : n_u + 1 \leq i \leq n_u + n_s\} < 0, \end{aligned} \quad \text{and} \quad r_n = 0,$$

where the integers n_u and n_s satisfy $n_u \geq 1$, $n_s \geq 1$, and $n_u + n_s + 1 = n$.

We will prove the following theorem about the persistence of T_0 in the full system (4.20) for small ϵ .

Theorem 4.4. *Consider (4.20) with μ , ν , and γ satisfying (4.21). Suppose Assumption 4.2 and Assumption 4.3 hold. Then there exists an $\epsilon^* \in (0, \epsilon_0]$ such that for (4.20) with any $\epsilon \in (0, \epsilon^*]$, there is a unique invariant torus T_ϵ inside an $\mathcal{O}(1)$ -neighborhood (i.e., independent of ϵ) of T_0 . In addition, T_ϵ is C^r diffeomorphic and $\mathcal{O}(\epsilon^\mu)$ -close to T_0 . In particular, T_ϵ has the parameterization*

$$(4.23) \quad T_\epsilon = \{(w, \theta) : w = \chi(\zeta) + \hat{n}(\zeta) \varrho(\zeta, \theta; \epsilon), \zeta \in \mathbb{R}(\text{mod } \zeta_0), \theta \in \mathbb{T}^m\},$$

where $\varrho(\cdot, \cdot; \epsilon) : \mathbb{R}(\text{mod } \zeta_0) \times \mathbb{T}^m \rightarrow \mathbb{R}^{n-1}$ is an $\mathcal{O}(\epsilon^\mu)$, C^r function and $\hat{n} : \mathbb{R}(\text{mod } \zeta_0) \rightarrow \mathbb{R}^{n \times (n-1)}$ is a C^{r+1} matrix function such that for each $\zeta \in \mathbb{R}(\text{mod } \zeta_0)$ the columns of $\hat{n}(\zeta)$ form a basis of the normal space of the periodic orbit S in \mathbb{R}^n at $\chi(\zeta)$.

Let $\eta : \mathbb{R} \rightarrow \mathbb{R}^{n \times (n-1)}$ be a C^{r+1} , bounded matrix function such that the $n \times n$ matrix $(\eta(\zeta) \frac{d}{d\zeta} \chi(\zeta))$ is nonsingular and its inverse is bounded for all $\zeta \in \mathbb{R}$. For a sufficiently small $\Delta > 0$, we make a C^{r+1} change of variables in a small neighborhood of the periodic orbit S as follows:

$$(4.24) \quad w = \chi(\zeta) + \eta(\zeta) \begin{pmatrix} a \\ b \end{pmatrix},$$

where $\zeta \in \mathbb{R}$, $a \in \mathbb{R}^{n_u}$ with $\|a\| < \Delta$, and $b \in \mathbb{R}^{n_s}$ with $\|b\| < \Delta$. By Floquet's theorem, Assumption 4.3 implies that we can choose a $2\zeta_0$ -periodic $\eta(\zeta)$ such that in the a - b - ζ coordinates, the averaged equation $\dot{w} = \epsilon F_1(w)$ is transformed into the following normal form:

$$(4.25) \quad \begin{aligned} \dot{a} &= \epsilon(A_0 a + A_1(a, b, \zeta)), \\ \dot{b} &= \epsilon(B_0 b + B_1(a, b, \zeta)), \\ \dot{\zeta} &= \epsilon(1 + V_1(a, b, \zeta)), \end{aligned}$$

where both the $n_u \times n_u$ constant matrix A_0 and the $n_s \times n_s$ constant matrix B_0 are in real Jordan form, and the functions $A_1(a, b, \zeta)$, $B_1(a, b, \zeta)$, and $V_1(a, b, \zeta)$ are all $\mathcal{O}(\|a\|^2 + \|b\|^2)$ and $2\zeta_0$ -periodic in ζ . Furthermore, the real parts of the eigenvalues of A_0 coincide with the Lyapunov exponents r_1, \dots, r_{n_u} , and the real parts of the eigenvalues of B_0 coincide with the Lyapunov exponents $r_{n_u+1}, \dots, r_{n_u+n_s}$.

We will make a further change of coordinates by rescaling individual components of a and b so that we can verify (1.6a) for the transformed system. Since A_0 and B_0 are in real Jordan form, it suffices to illustrate how to rescale the components of a that are associated with the same Jordan block $A_{0,j}$ of A_0 . Suppose we have

$$A_{0,j} = \begin{pmatrix} r_\kappa & \xi_\kappa & 1 & 0 & 0 & 0 \\ -\xi_\kappa & r_\kappa & 0 & 1 & 0 & 0 \\ 0 & 0 & r_\kappa & \xi_\kappa & 1 & 0 \\ 0 & 0 & -\xi_\kappa & r_\kappa & 0 & 1 \\ 0 & 0 & 0 & 0 & r_\kappa & \xi_\kappa \\ 0 & 0 & 0 & 0 & -\xi_\kappa & r_\kappa \end{pmatrix},$$

where r_κ is one of the positive Lyapunov exponents specified in Assumption 4.3, and $r_\kappa \pm \xi_\kappa i$ are the pair of complex eigenvalues of $A_{0,j}$. Let $a_{j,1}, \dots, a_{j,6}$ be the

corresponding components of a that are associated with $A_{0,j}$. Define

$$\begin{pmatrix} p_{j,1} \\ \vdots \\ p_{j,6} \end{pmatrix} := \Lambda_j \begin{pmatrix} a_{j,1} \\ \vdots \\ a_{j,6} \end{pmatrix}, \quad \text{where } \Lambda_j^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{j,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{j,1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{j,2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{j,2} \end{pmatrix}.$$

It follows that

$$P_{0,j} := \Lambda_j A_{0,j} \Lambda_j^{-1} = \begin{pmatrix} r_\kappa & \xi_\kappa & \lambda_{j,1} & 0 & 0 & 0 \\ -\xi_\kappa & r_\kappa & 0 & \lambda_{j,1} & 0 & 0 \\ 0 & 0 & r_\kappa & \xi_\kappa & \frac{\lambda_{j,2}}{\lambda_{j,1}} & 0 \\ 0 & 0 & -\xi_\kappa & r_\kappa & 0 & \frac{\lambda_{j,2}}{\lambda_{j,1}} \\ 0 & 0 & 0 & 0 & r_\kappa & \xi_\kappa \\ 0 & 0 & 0 & 0 & -\xi_\kappa & r_\kappa \end{pmatrix}.$$

Notice that for any $c > 0$, there exist $\lambda_{j,1}, \lambda_{j,2} > 0$ such that

$$\langle y, P_{0,j} y \rangle \geq (r_\kappa - c) \|y\|^2 \quad \text{for any } y \in \mathbb{R}^6.$$

For each Jordan block of A_0 and B_0 , we apply similar rescaling if necessary to the corresponding components of a and b so that under the change of coordinates $(a, b) \mapsto (p, q)$, (4.25) becomes

$$\begin{aligned} \dot{p} &= \epsilon(P_0 p + P_1(p, q, \zeta)), \\ \dot{q} &= \epsilon(Q_0 q + Q_1(p, q, \zeta)), \\ \dot{\zeta} &= \epsilon(1 + Z_1(p, q, \zeta)), \end{aligned}$$

with P_0 and Q_0 now satisfying

$$(4.26a) \quad \langle p, P_0 p \rangle \geq \sigma \|p\|^2 \quad \text{for any } p \in \mathbb{R}^{n_u},$$

$$(4.26b) \quad \langle q, Q_0 q \rangle \leq -\sigma \|q\|^2 \quad \text{for any } q \in \mathbb{R}^{n_s},$$

where σ is a positive constant satisfying $0 < \sigma < \min\{\sigma_s, \sigma_u\}$. We denote this change of coordinates by $(a, b) = v(p, q)$.

We now obtain a normal form of the full system (4.20) near the torus T_0 in terms of (p, q, ζ, θ) as follows:

$$(4.27) \quad \begin{aligned} \dot{p} &= \epsilon(P_0 p + P_1(p, q, \zeta)) + \epsilon^{1+\mu} P_2(p, q, \zeta, \theta, \epsilon), \\ \dot{q} &= \epsilon(Q_0 q + Q_1(p, q, \zeta)) + \epsilon^{1+\mu} Q_2(p, q, \zeta, \theta, \epsilon), \\ \dot{\zeta} &= \epsilon(1 + Z_1(p, q, \zeta)) + \epsilon^{1+\mu} Z_2(p, q, \zeta, \theta, \epsilon), \\ \dot{\theta} &= \Theta_0 + \epsilon^\nu \Theta_1(p, q, \zeta, \epsilon) + \epsilon^{1+\gamma} \Theta_2(p, q, \zeta, \theta, \epsilon), \end{aligned}$$

where $\|p\| < \delta_0$ and $\|q\| < \delta_0$ for a certain $\delta_0 > 0$ such that $v(p, q) \in \{(a, b) : \|a\| < \Delta, \|b\| < \Delta\}$. In addition to taking $\zeta \in \mathbb{R}$, we lift θ to \mathbb{R}^m in the subsequent analysis. Based on Assumptions 4.2 and 4.3 and the preceding changes of variables, we can easily verify a set of properties regarding the smoothness, boundedness, and periodicity of the functions on the right-hand side of (4.27). We state these properties in the following lemma while omitting their straightforward verifications.

Lemma 4.5. *Let $D_0 := \{(p, q) : \|p\| < \delta_0, \|q\| < \delta_0\}$.*

- (1) The constant matrices P_0 and Q_0 satisfy (4.26).
- (2) The functions P_1 , Q_1 , and Z_1 are C^r on $D_0 \times \mathbb{R}$.
- (3) The function Θ_1 is C^r with respect to (p, q, ζ) and continuous with respect to ϵ on $D_0 \times \mathbb{R} \times [0, \epsilon_0]$.
- (4) The functions P_2 , Q_2 , Z_2 , and Θ_2 are C^r with respect to (p, q, ζ, θ) and continuous with respect to ϵ on $D_0 \times \mathbb{R} \times \mathbb{R}^m \times [0, \epsilon_0]$.
- (5) There exist positive constants C_1 and C_2 such that for any $\delta \in [0, \delta_0]$ and any $\|p\| \leq \delta$, $\|q\| \leq \delta$, and $\zeta \in \mathbb{R}$,

$$\begin{aligned} \|P_1\| &\leq C_1 \delta^2, & \|Q_1\| &\leq C_1 \delta^2, & \|Z_1\| &\leq C_1 \delta^2, \\ \|D_\zeta P_1\| &\leq C_1 \delta^2, & \|D_\zeta Q_1\| &\leq C_1 \delta^2, & \|D_\zeta Z_1\| &\leq C_1 \delta^2, \\ \|D_\xi P_1\| &\leq C_2 \delta, & \|D_\xi Q_1\| &\leq C_2 \delta, & \|D_\xi Z_1\| &\leq C_2 \delta, \end{aligned}$$

where $\xi = p$ or q in the last three inequalities.

- (6) There exist positive constants C_3 and C_4 such that for any $(p, q) \in D_0$, $\zeta \in \mathbb{R}$, $\theta \in \mathbb{R}^m$, and $\epsilon \in [0, \epsilon_0]$,

$$\begin{aligned} \|P_2\| &\leq C_3, & \|Q_2\| &\leq C_3, & \|Z_2\| &\leq C_3, \\ \|\Theta_1\| &\leq C_3, & \|\Theta_2\| &\leq C_3, \\ \|D_\xi P_2\| &\leq C_4, & \|D_\xi Q_2\| &\leq C_4, & \|D_\xi Z_2\| &\leq C_4, \\ \|D_\xi \Theta_1\| &\leq C_4, & \|D_\xi \Theta_2\| &\leq C_4, \end{aligned}$$

where $\xi = p, q, \zeta$, or θ .

- (7) The second to r -th derivatives of P_1 , P_2 , Q_1 , Q_2 , Z_1 , Z_2 , Θ_1 , and Θ_2 with respect to (p, q, ζ) or (p, q, ζ, θ) are all bounded on the corresponding domains specified in (2), (3), and (4).
- (8) The functions P_1 , P_2 , Q_1 , Q_2 , Z_1 , Z_2 , Θ_1 , and Θ_2 are $2\zeta_0$ -periodic in ζ and 2π -periodic in each component of θ on the corresponding domains specified in (2), (3), and (4).

We now show that for every fixed, sufficiently small $\epsilon > 0$, (4.27) has a C^r positively invariant manifold. We introduce another two rescaled variables: $\bar{q} := q/2$ and $\bar{\theta} := k^{-1}\theta$ with the scaling factor $k > 0$ to be determined later. By taking $z = (\bar{q}, \zeta, \bar{\theta})$ and treating p as “ a ”, we organize the rescaled system into the form of (1.3) as follows:

$$(4.28) \quad \begin{aligned} \dot{p} &= f(p, z) := \epsilon(P_0 p + P_1(p, 2\bar{q}, \zeta)) + \epsilon^{1+\mu} P_2(p, 2\bar{q}, \zeta, k\bar{\theta}, \epsilon), \\ \dot{z} &= g(p, z) := \begin{pmatrix} \epsilon(Q_0 \bar{q} + \frac{1}{2} Q_1(p, 2\bar{q}, \zeta)) + \frac{1}{2} \epsilon^{1+\mu} Q_2(p, 2\bar{q}, \zeta, k\bar{\theta}, \epsilon) \\ \epsilon(1 + Z_1(p, 2\bar{q}, \zeta)) + \epsilon^{1+\mu} Z_2(p, 2\bar{q}, \zeta, k\bar{\theta}, \epsilon) \\ \frac{1}{k} \Theta_0 + \frac{1}{k} \epsilon^\nu \Theta_1(p, 2\bar{q}, \zeta, \epsilon) + \frac{1}{k} \epsilon^{1+\gamma} \Theta_2(p, 2\bar{q}, \zeta, k\bar{\theta}, \epsilon) \end{pmatrix}. \end{aligned}$$

We shall restrict $(p, z) = (p, \bar{q}, \zeta, \bar{\theta})$ to the domain

$$U_\delta := \{\|p\| < \delta\} \times \{\|\bar{q}\| < \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}^m,$$

where $\delta \in (0, \delta_0]$ is to be determined. Obviously, U_δ satisfies Hypothesis 1. As we have done for the previous two examples, we define $\Gamma_{\epsilon, k, \delta} \subset U_\delta$ to be the set of points that stay in U_δ forever in forward time along solution trajectories of (4.28), that is,

$$(4.29) \quad \Gamma_{\epsilon, k, \delta} := \{(p, z) : \Phi_{\epsilon, k}(t, (p, z)) \in U_\delta \text{ for all } t \geq 0\},$$

where $\Phi_{\epsilon,k}(t, (p, z))$ is the flow of (4.28). By this definition, $\Gamma_{\epsilon,k,\delta}$ is the largest positively invariant subset of U_δ under the flow $\Phi_{\epsilon,k}$. Then we need to prove that $\Gamma_{\epsilon,k,\delta}$ satisfies

$$(4.30) \quad \Pi_\perp(\Gamma_{\epsilon,k,\delta}) = \Pi_\perp(U_\delta) = \{\|\bar{q}\| < \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}^m,$$

where Π_\perp is the projection onto the z -coordinate. Again, we will achieve this using the Ważewski theorem. However, since p sits in \mathbb{R}^{n_u} , some arguments are slightly different.

Lemma 4.6. *There exist $\epsilon_1 \in (0, \epsilon_0]$ and $\delta_1 \in (0, \delta_0]$ satisfying $2\frac{C_3}{\sigma}\epsilon_1^\mu < \delta_1$ such that for (4.28) with any $\epsilon \in (0, \epsilon_1]$ and U_δ with any $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta_1]$, the positively invariant set $\Gamma_{\epsilon,k,\delta} \subset U_\delta$ satisfies (4.30).*

Proof. We partition the boundary of U_δ into two subsets S_1 and S_2 as follows:

$$\begin{aligned} S_1 &:= \{\|p\| = \delta\} \times \{\|\bar{q}\| \leq \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}^m, \\ S_2 &:= \{\|p\| < \delta\} \times \{\|\bar{q}\| = \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}^m. \end{aligned}$$

Using (4.26a) and the estimates in part (5) and part (6) of Lemma 4.5, we obtain that for any point in S_1 ,

$$\begin{aligned} \langle p, \dot{p} \rangle &= \langle p, \epsilon(P_0 p + P_1(p, 2\bar{q}, \zeta)) \rangle + \langle p, \epsilon^{1+\mu} P_2(p, 2\bar{q}, \zeta, k\bar{\theta}, \epsilon) \rangle \\ &\geq \epsilon \langle p, P_0 p \rangle - \epsilon C_1 \delta^2 \|p\| - \epsilon^{1+\mu} C_3 \|p\| \\ &\geq \epsilon \sigma \|p\|^2 - \epsilon C_1 \delta^2 \|p\| - \epsilon^{1+\mu} C_3 \|p\| \\ &= \epsilon \delta (\sigma \delta - C_1 \delta^2 - \epsilon^\mu C_3). \end{aligned}$$

Similarly, for points in S_2 , we have

$$\begin{aligned} \langle \bar{q}, \dot{\bar{q}} \rangle &= \langle \bar{q}, \epsilon(Q_0 \bar{q} + \frac{1}{2} Q_1(p, 2\bar{q}, \zeta)) \rangle + \langle \bar{q}, \frac{1}{2} \epsilon^{1+\mu} Q_2(p, 2\bar{q}, \zeta, k\bar{\theta}, \epsilon) \rangle \\ &\leq \epsilon \langle \bar{q}, Q_0 \bar{q} \rangle + \frac{1}{2} \epsilon C_1 \delta^2 \|\bar{q}\| + \frac{1}{2} \epsilon^{1+\mu} C_3 \|\bar{q}\| \\ &\leq -\epsilon \sigma \|\bar{q}\|^2 + \frac{1}{2} \epsilon C_1 \delta^2 \|\bar{q}\| + \frac{1}{2} \epsilon^{1+\mu} C_3 \|\bar{q}\| \\ &= -\epsilon \frac{\delta}{4} (\sigma \delta - C_1 \delta^2 - \epsilon^\mu C_3). \end{aligned}$$

Since for $\mu > 0$, $\epsilon^\mu \rightarrow 0$ as $\epsilon \rightarrow 0^+$, we can take $\epsilon_1 \in (0, \epsilon_0]$ and $\delta_1 \in (0, \delta_0]$ such that $2\frac{C_3}{\sigma}\epsilon_1^\mu < \delta_1 \leq \frac{\sigma}{2C_1}$. Then for any $\epsilon \in (0, \epsilon_1]$ and any $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta_1]$,

$$\begin{aligned} \sigma \delta - C_1 \delta^2 - \epsilon^\mu C_3 &\geq \sigma(2\frac{C_3}{\sigma}\epsilon^\mu) - C_1(2\frac{C_3}{\sigma}\epsilon^\mu)^2 - \epsilon^\mu C_3 \\ &= C_3 \epsilon^\mu - 4C_1 C_3^2 \frac{1}{\sigma^2} \epsilon^{2\mu} > 0, \end{aligned}$$

where the last inequality can be guaranteed by choosing ϵ_1 sufficiently small. It follows that $\frac{d}{dt}\|p\|^2 > 0$ for any point in S_1 and $\frac{d}{dt}\|\bar{q}\|^2 < 0$ for any point in S_2 . Then solution trajectories of (4.28) leave and enter $\text{cl}(U_\delta)$ through points in S_1 and S_2 , respectively. In particular, S_1 is the set of points through which trajectories leave $\text{cl}(U_\delta)$ immediately in forward time. Then it can be easily verified that $\text{cl}(U_\delta)$ is a Ważewski set. Let $U_\delta^0 \subseteq \text{cl}(U_\delta)$ be the set of points that do not stay in $\text{cl}(U_\delta)$ forever in forward time. By the Ważewski theorem, there exists a continuous function $\mathcal{R} : U_\delta^0 \times [0, 1] \rightarrow U_\delta^0$ such that \mathcal{R} is a strong deformation retraction of U_δ^0 onto S_1 .

Suppose that for a certain $z^* = (\bar{q}^*, \zeta^*, \bar{\theta}^*) \in \{\|\bar{q}\| < \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}$,

$$(4.31) \quad \Gamma_{\epsilon,k,\delta} \cap \{(p, z^*) : \|p\| < \delta\} = \emptyset,$$

where the second set on the left is just the section of U_δ at z^* . Since $\{(p, z^*) : \|p\| = \delta\} \subset S_1 \subset U_\delta^0$, (4.31) implies that the set $\{(p, z^*) : \|p\| \leq \delta\}$ is contained inside the domain U_δ^0 of the continuous function $\mathcal{R}(\cdot, 1) : U_\delta^0 \rightarrow S_1$. This allows us to construct a continuous function that maps the closed n_u -ball $\{\|p\| \leq \delta\}$ to its boundary $\{\|p\| = \delta\}$:

$$p \mapsto \Pi \circ \mathcal{R}((p, z^*), 1),$$

where Π is the projection onto the p -coordinate. Since \mathcal{R} is a strong deformation retraction of U_δ^0 onto S_1 , $\mathcal{R}((p, z^*), 1) = (p, z^*)$ for any $(p, z^*) \in S_1$. Thus, the above continuous function has the property that for any p with $\|p\| = \delta$,

$$\Pi \circ \mathcal{R}((p, z^*), 1) = \Pi(p, z^*) = p.$$

The existence of such a continuous function contradicts the fact that there is no retraction that maps a closed n -ball onto its boundary (i.e., an $(n-1)$ -sphere). Thus, $\Gamma_{\epsilon, k, \delta} \cap \{(p, z) : \|p\| < \delta\} \neq \emptyset$ for any $z \in \{\|\bar{q}\| < \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}^m$. This proves (4.30). \square

Next, using (1), (5), and (6) of Lemma 4.5, we obtain the following estimates regarding $D_p f(p, z)$, $D_z g(p, z)$, $D_z f(p, z)$, and $D_p g(p, z)$ for any $(p, z) \in U_\delta$.

(1) For any $p' \in \mathbb{R}^{n_u}$,

$$\begin{aligned} \langle p', D_p f p' \rangle &= \langle p', (\epsilon P_0 + \epsilon D_p P_1 + \epsilon^{1+\mu} D_p P_2) p' \rangle \\ &\geq \epsilon \sigma \|p'\|^2 - \epsilon C_2 \delta \|p'\|^2 - \epsilon^{1+\mu} C_4 \|p'\|^2 \\ &= \alpha_u \|p'\|^2, \end{aligned}$$

where

$$\alpha_u := \epsilon(\sigma - C_2 \delta) - \epsilon^{1+\mu} C_4.$$

(2) For any $z' = (\bar{q}', \zeta', \bar{\theta}') \in \mathbb{R}^{n_s} \times \mathbb{R} \times \mathbb{R}^m$,

$$\begin{aligned} \langle z', D_z g z' \rangle &= \langle \bar{q}', (\epsilon Q_0 + \epsilon D_q Q_1 + \epsilon^{1+\mu} D_q Q_2) \bar{q}' \rangle + \langle \bar{q}', \frac{1}{2} k \epsilon^{1+\mu} D_\theta Q_2 \bar{\theta}' \rangle \\ &\quad + \langle \bar{q}', (\frac{1}{2} \epsilon D_\zeta Q_1 + \frac{1}{2} \epsilon^{1+\mu} D_\zeta Q_2) \zeta' \rangle \\ &\quad + \langle \zeta', (2\epsilon D_q Z_1 + 2\epsilon^{1+\mu} D_q Z_2) \bar{q}' \rangle + \langle \zeta', k \epsilon^{1+\mu} D_\theta Z_2 \bar{\theta}' \rangle \\ &\quad + \langle \zeta', (\epsilon D_\zeta Z_1 + \epsilon^{1+\mu} D_\zeta Z_2) \zeta' \rangle \\ &\quad + \langle \bar{\theta}', (2\frac{1}{k} \epsilon^\nu D_q \Theta_1 + 2\frac{1}{k} \epsilon^{1+\gamma} D_q \Theta_2) \bar{q}' \rangle + \langle \bar{\theta}', \epsilon^{1+\gamma} D_\theta \Theta_2 \bar{\theta}' \rangle \\ &\quad + \langle \bar{\theta}', (\frac{1}{k} \epsilon^\nu D_\zeta \Theta_1 + \frac{1}{k} \epsilon^{1+\gamma} D_\zeta \Theta_2) \zeta' \rangle \\ &\leq -\epsilon \sigma \|\bar{q}'\|^2 + (\epsilon C_2 \delta + \epsilon^{1+\mu} C_4) \|\bar{q}'\|^2 + \frac{1}{2} k \epsilon^{1+\mu} C_4 \|\bar{q}'\| \|\bar{\theta}'\| \\ &\quad + (\frac{1}{2} \epsilon C_1 \delta^2 + \frac{1}{2} \epsilon^{1+\mu} C_4) \|\bar{q}'\| |\zeta'| \\ &\quad + (2\epsilon C_2 \delta + 2\epsilon^{1+\mu} C_4) |\zeta'| \|\bar{q}'\| + k \epsilon^{1+\mu} C_4 |\zeta'| \|\bar{\theta}'\| \\ &\quad + (\epsilon C_1 \delta^2 + \epsilon^{1+\mu} C_4) |\zeta'|^2 \\ &\quad + (2\frac{1}{k} \epsilon^\nu C_4 + 2\frac{1}{k} \epsilon^{1+\gamma} C_4) \|\bar{\theta}'\| \|\bar{q}'\| + \epsilon^{1+\gamma} C_4 \|\bar{\theta}'\|^2 \\ &\quad + (\frac{1}{k} \epsilon^\nu C_4 + \frac{1}{k} \epsilon^{1+\gamma} C_4) \|\bar{\theta}'\| |\zeta'| \\ &\leq \ell_u \|z'\|^2, \end{aligned}$$

where

$$\begin{aligned} \ell_u &:= \epsilon(3C_2 \delta + \frac{3}{2} C_1 \delta^2) + \epsilon^{1+\mu} \frac{9}{2} C_4 + \epsilon^{1+\gamma} C_4 \\ &\quad + k \epsilon^{1+\mu} \frac{3}{2} C_4 + \frac{1}{k} (\epsilon^\nu 3C_4 + \epsilon^{1+\gamma} 3C_4). \end{aligned}$$

To obtain the last inequality, we have discarded the negative term $-\epsilon\sigma\|\bar{q}'\|^2$ and then used $\|\bar{q}'\| \leq \|z'\|$, $|\zeta'| \leq \|z'\|$, and $\|\bar{\theta}'\| \leq \|z'\|$.

(3)

$$\begin{aligned} \|D_z f\| &\leq (\|2\epsilon D_q P_1\| + \|2\epsilon^{1+\mu} D_q P_2\|) + (\|\epsilon D_\zeta P_1\| + \|\epsilon^{1+\mu} D_\zeta P_2\|) \\ &\quad + \|k\epsilon^{1+\mu} D_\theta P_2\| \\ &\leq (2\epsilon C_2 \delta + 2\epsilon^{1+\mu} C_4) + (\epsilon C_1 \delta^2 + \epsilon^{1+\mu} C_4) + k\epsilon^{1+\mu} C_4 \\ &= L_{f,z}, \end{aligned}$$

where

$$L_{f,z} := \epsilon(2C_2\delta + C_1\delta^2) + \epsilon^{1+\mu}3C_4 + k\epsilon^{1+\mu}C_4.$$

(4)

$$\begin{aligned} \|D_p g\| &\leq (\|\tfrac{1}{2}\epsilon D_p Q_1\| + \|\tfrac{1}{2}\epsilon^{1+\mu} D_p Q_2\|) + (\|\epsilon D_p Z_1\| + \|\epsilon^{1+\mu} D_p Z_2\|) \\ &\quad + (\|\tfrac{1}{k}\epsilon^\nu D_p \Theta_1\| + \|\tfrac{1}{k}\epsilon^{1+\gamma} D_p \Theta_2\|) \\ &\leq (\tfrac{1}{2}\epsilon C_2 \delta + \tfrac{1}{2}\epsilon^{1+\mu} C_4) + (\epsilon C_2 \delta + \epsilon^{1+\mu} C_4) + (\tfrac{1}{k}\epsilon^\nu C_4 + \tfrac{1}{k}\epsilon^{1+\gamma} C_4) \\ &= L_{g,p}, \end{aligned}$$

where

$$L_{g,p} := \epsilon \tfrac{3}{2} C_2 \delta + \epsilon^{1+\mu} \tfrac{3}{2} C_4 + \tfrac{1}{k} (\epsilon^\nu C_4 + \epsilon^{1+\gamma} C_4).$$

In order to verify Hypothesis 2 and Hypothesis 2* for (4.28) on U_δ , we define an auxiliary function $\mathcal{K} : (0, \epsilon_0] \times (0, \infty) \times (0, \delta_0] \rightarrow \mathbb{R}$ as follows:

$$\begin{aligned} \mathcal{K}(\epsilon, k, \delta) &:= \alpha_u - (r\ell_u + (r+1)L_{g,p} + L_{f,z}) \\ &= \epsilon\sigma - \epsilon(K_1\delta + K_2\delta^2 + \epsilon^\mu K_3 + \epsilon^\gamma K_4 + k\epsilon^\mu K_5 + \tfrac{1}{k}(\epsilon^{\nu-1}K_6 + \epsilon^\gamma K_7)), \end{aligned}$$

where r is the degree of smoothness referred to in (2) and (3) of Lemma 4.5 and the constants K_i are defined as follows:

$$\begin{aligned} K_1 &:= \tfrac{9}{2}(r+1)C_2, \\ K_2 &:= (\tfrac{3}{2}r+1)C_1, \\ K_3 &:= (6r + \tfrac{11}{2})C_4, \\ K_4 &:= rC_4, \\ K_5 &:= (\tfrac{3}{2}r+1)C_4, \\ K_6 &:= (4r+1)C_4, \\ K_7 &:= (4r+1)C_4. \end{aligned}$$

Lemma 4.7. *There exist $\epsilon_2 \in (0, \epsilon_0]$ and $\delta_2 \in (0, \delta_0]$ satisfying $2\frac{C_3}{\sigma}\epsilon_2^\mu < \delta_2$ such that $\mathcal{K}(\epsilon, k_\epsilon, \delta) > 0$ for any $\epsilon \in (0, \epsilon_2]$, any $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta_2]$, and an appropriately chosen k_ϵ that depends on ϵ .*

Proof. Note that for any $\epsilon > 0$,

$$\inf_{k \in (0, \infty)} \left\{ k\epsilon^\mu K_5 + \tfrac{1}{k}(\epsilon^{\nu-1}K_6 + \epsilon^\gamma K_7) \right\} = (\epsilon^{\mu+\nu-1}K_5K_6 + \epsilon^{\mu+\gamma}K_5K_7)^{\frac{1}{2}}.$$

In particular, for every $\epsilon > 0$, there exists a $k_\epsilon > 0$ such that

$$(4.32) \quad k_\epsilon \epsilon^\mu K_5 + \tfrac{1}{k_\epsilon}(\epsilon^{\nu-1}K_6 + \epsilon^\gamma K_7) \leq (\epsilon^{\mu+\nu-1}K_5K_6 + \epsilon^{\mu+\gamma}K_5K_7)^{\frac{1}{2}} + \tfrac{1}{3}\sigma.$$

By (4.21), we have $\epsilon^\mu \rightarrow 0$, $\epsilon^\gamma \rightarrow 0$, $\epsilon^{\mu+\gamma} \rightarrow 0$, and $\epsilon^{\mu+\nu-1} \rightarrow 0$ for $\epsilon \rightarrow 0^+$. Thus, there exist $\epsilon_2 \in (0, \epsilon_0]$ and $\delta_2 \in (0, \delta_0]$ satisfying $2\frac{C_3}{\sigma}\epsilon_2^\mu < \delta_2$ such that for any $\epsilon \in (0, \epsilon_2]$ and any $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta_2]$,

$$\begin{aligned} \frac{2}{3}\sigma &\geq K_1\delta + K_2\delta^2 + \epsilon^\mu K_3 + \epsilon^\gamma K_4 + (\epsilon^{\mu+\nu-1}K_5K_6 + \epsilon^{\mu+\gamma}K_5K_7)^{\frac{1}{2}} + \frac{1}{3}\sigma \\ &\geq K_1\delta + K_2\delta^2 + \epsilon^\mu K_3 + \epsilon^\gamma K_4 + k_\epsilon\epsilon^\mu K_5 + \frac{1}{k_\epsilon}(\epsilon^{\nu-1}K_6 + \epsilon^\gamma K_7). \end{aligned}$$

It follows that $\mathcal{K}(\epsilon, k_\epsilon, \delta) \geq \frac{1}{3}\epsilon\sigma > 0$. \square

Lemma 4.7 guarantees that both (1.6c) and (1.6d) hold for (4.28) on U_δ . Thus, by combining Lemma 4.6 and Lemma 4.7, we establish the existence and the C^r smoothness of a positively invariant manifold for (4.28) as stated in the following lemma.

Lemma 4.8. *There exist $\epsilon^* \in (0, \epsilon_0]$ and $\delta^* \in (0, \delta_0]$ satisfying $2\frac{C_3}{\sigma}\epsilon^{*\mu} < \delta^*$ such that for (4.28) with any $\epsilon \in (0, \epsilon^*]$ and any $k = k_\epsilon$ that satisfies (4.32), the positively invariant set $\Gamma_{\epsilon, k_\epsilon, \delta^*} \subset U_{\delta^*}$ is the graph of a C^r function $h(\cdot, \cdot, \cdot; \epsilon, k_\epsilon) : \Pi_\perp(U_{\delta^*}) \rightarrow \{\|p\| < \delta^*\}$ that satisfies*

$$(4.33) \quad \|h(\bar{q}_2, \zeta_2, \bar{\theta}_2; \epsilon, k_\epsilon) - h(\bar{q}_1, \zeta_1, \bar{\theta}_1; \epsilon, k_\epsilon)\| \leq \|\bar{q}_2 - \bar{q}_1\| + |\zeta_2 - \zeta_1| + \|\bar{\theta}_2 - \bar{\theta}_1\|$$

for any $(\bar{q}_1, \zeta_1, \bar{\theta}_1)$ and $(\bar{q}_2, \zeta_2, \bar{\theta}_2) \in \Pi_\perp(U_{\delta^*})$. Furthermore, for each $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta^*]$, the positively invariant set $\Gamma_{\epsilon, k_\epsilon, \delta} = \Gamma_{\epsilon, k_\epsilon, \delta^*} \cap U_\delta$.

Proof. Take $\epsilon^* = \min\{\epsilon_1, \epsilon_2\}$ and $\delta^* = \min\{\delta_1, \delta_2\}$. Clearly, $2\frac{C_3}{\sigma}\epsilon^{*\mu} < \delta^*$. Thus, for (4.28) with any $\epsilon \in (0, \epsilon^*]$ and any $k = k_\epsilon$ that satisfies (4.32), Lemma 4.6 and Lemma 4.7 guarantee that both Theorem 1.1 and Theorem 1.3 are applicable on the domain $U_\delta = \{\|p\| < \delta\} \times \{\|\bar{q}\| < \frac{\delta}{2}\} \times \mathbb{R} \times \mathbb{R}^m$ with any $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta^*]$. In particular, for $\delta = \delta^*$, the positively invariant set $\Gamma_{\epsilon, k_\epsilon, \delta^*} \subset U_{\delta^*}$ is the graph of a C^r function

$$h(\cdot, \cdot, \cdot; \epsilon, k_\epsilon) : \Pi_\perp(U_{\delta^*}) = \{\|\bar{q}\| < \frac{\delta^*}{2}\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|p\| < \delta^*\}$$

that satisfies (4.33) according to Theorem 1.1 or Theorem 1.3. Furthermore, for each $\delta \in [2\frac{C_3}{\sigma}\epsilon^\mu, \delta^*]$, $\Pi_\perp(\Gamma_{\epsilon, k_\epsilon, \delta}) = \Pi_\perp(U_\delta)$ by Lemma 4.6, and $\Gamma_{\epsilon, k_\epsilon, \delta} \subseteq \Gamma_{\epsilon, k_\epsilon, \delta^*} \cap U_\delta$ by the definition (4.29). Thus $\Gamma_{\epsilon, k_\epsilon, \delta} = \Gamma_{\epsilon, k_\epsilon, \delta^*} \cap U_\delta$. \square

We also have the following facts regarding the function h .

- (1) For any $\epsilon \in (0, \epsilon^*]$ and any k_ϵ and k'_ϵ that both satisfy (4.32),

$$(4.34) \quad h(\bar{q}, \zeta, \frac{\theta}{k_\epsilon}; \epsilon, k_\epsilon) = h(\bar{q}, \zeta, \frac{\theta}{k'_\epsilon}; \epsilon, k'_\epsilon)$$

for all $(\bar{q}, \zeta, \bar{\theta}) \in \{\|\bar{q}\| < \frac{\delta^*}{2}\} \times \mathbb{R} \times \mathbb{R}^m$.

- (2) For any $\epsilon \in (0, \epsilon^*]$ and any k_ϵ that satisfies (4.32), $h(\bar{q}, \zeta, \bar{\theta}; \epsilon, k_\epsilon)$ is $2\zeta_0$ -periodic in ζ and $\frac{2\pi}{k_\epsilon}$ -periodic in each component of $\bar{\theta}$.

- (3) For any $\epsilon \in (0, \epsilon^*]$ and any k_ϵ that satisfies (4.32),

$$\|h(\bar{q}, \zeta, \bar{\theta}; \epsilon, k_\epsilon)\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu$$

for all $(\bar{q}, \zeta, \bar{\theta}) \in \{\|\bar{q}\| < \frac{C_3}{\sigma}\epsilon^\mu\} \times \mathbb{R} \times \mathbb{R}^m$.

The first two properties should be familiar by now, and they follow from the definition (4.29) and the same arguments as given in Subsections 4.1 and 4.2 for the first two examples. The third property is a simple consequence of the fact that $\Gamma_{\epsilon, k_\epsilon, \delta} = \Gamma_{\epsilon, k_\epsilon, \delta^*} \cap U_\delta$ for $\delta = 2\frac{C_3}{\sigma}\epsilon^\mu$.

By (4.34), we can define a C^r function $\mathcal{P}(\cdot, \cdot, \cdot; \epsilon) : \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|p\| < \delta^*\}$ for each $\epsilon \in (0, \epsilon^*]$ as follows:

$$\mathcal{P}(q, \zeta, \theta; \epsilon) := h\left(\frac{q}{2}, \zeta, \frac{\theta}{k_\epsilon}; \epsilon, k_\epsilon\right),$$

where k_ϵ can be taken to be any value that satisfies (4.32). Then it follows immediately from (4.33) that

$$(4.35) \quad \|\mathcal{P}(q_2, \zeta_2, \theta_2; \epsilon) - \mathcal{P}(q_1, \zeta_1, \theta_1; \epsilon)\| \leq \frac{1}{2}\|q_2 - q_1\| + |\zeta_2 - \zeta_1| + \frac{1}{k_\epsilon}\|\theta_2 - \theta_1\|$$

for any (q_1, ζ_1, θ_1) and $(q_2, \zeta_2, \theta_2) \in \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$. In addition, by the properties (2) and (3) above, $\mathcal{P}(q, \zeta, \theta; \epsilon)$ is $2\zeta_0$ -periodic in ζ and 2π -periodic in each component of θ , and

$$(4.36) \quad \|\mathcal{P}(q, \zeta, \theta; \epsilon)\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu$$

for all $(q, \zeta, \theta) \in \{\|q\| < 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \mathbb{R} \times \mathbb{R}^m$.

We now return to the normal form (4.27). The results stated in the next lemma are now obvious.

Lemma 4.9. *For any $\epsilon \in (0, \epsilon^*]$, (4.27) has a C^r positively invariant manifold W_ϵ^+ , which is the graph of the C^r function $\mathcal{P}(\cdot, \cdot, \cdot; \epsilon) : \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|p\| < \delta^*\}$. In addition, W_ϵ^+ is the largest positively invariant subset of $\{\|p\| < \delta^*\} \times \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$.*

Notice that all the analysis starting from the formulation of (4.28) up to Lemma 4.9 can be adapted (with only minor modifications) for the time reversal of (4.27). In particular, we let $\bar{p} := p/2$ while keeping $\bar{\theta} = k^{-1}\theta$, and then form a system in the spirit of (4.28) but using the time reversal of (4.27) and treating q and $(\bar{p}, \zeta, \bar{\theta})$ as “ a ” and “ z ” of (1.3), respectively. Then, by following all the previous steps, we construct a C^r function $\mathcal{Q}(\cdot, \cdot, \cdot; \epsilon) : \{\|p\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|q\| < \delta^*\}$ whose properties are completely analogous to those of the function $\mathcal{P}(\cdot, \cdot, \cdot; \epsilon)$, i.e.:

$$(4.37) \quad \|\mathcal{Q}(p_2, \zeta_2, \theta_2; \epsilon) - \mathcal{Q}(p_1, \zeta_1, \theta_1; \epsilon)\| \leq \frac{1}{2}\|p_2 - p_1\| + |\zeta_2 - \zeta_1| + \frac{1}{k_\epsilon}\|\theta_2 - \theta_1\|$$

for any (p_1, ζ_1, θ_1) and $(p_2, \zeta_2, \theta_2) \in \{\|p\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$; $\mathcal{Q}(p, \zeta, \theta; \epsilon)$ is $2\zeta_0$ -periodic in ζ and 2π -periodic in each component of θ ; and

$$(4.38) \quad \|\mathcal{Q}(p, \zeta, \theta; \epsilon)\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu$$

for all $(p, \zeta, \theta) \in \{\|p\| < 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \mathbb{R} \times \mathbb{R}^m$. Then we establish the existence and the C^r smoothness of a negatively invariant manifold for (4.27) as stated in the next lemma.

Lemma 4.10. *For any $\epsilon \in (0, \epsilon^*]$, (4.27) has a C^r negatively invariant manifold W_ϵ^- , which is the graph of the C^r function $\mathcal{Q}(\cdot, \cdot, \cdot; \epsilon) : \{\|p\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|q\| < \delta^*\}$. In addition, W_ϵ^- is the largest negatively invariant subset of $\{\|p\| < \delta^*\} \times \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$.*

Define $M_\epsilon := W_\epsilon^+ \cap W_\epsilon^-$ for each $\epsilon \in (0, \epsilon^*]$. Clearly, M_ϵ is invariant (i.e., in both forward time and backward time) under the flow of (4.27) with the corresponding ϵ . Below we show that M_ϵ is in fact the graph of a C^r function that maps $\mathbb{R} \times \mathbb{R}^m$ into $\{\|p\| < 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| < 2\frac{C_3}{\sigma}\epsilon^\mu\}$ and thus a C^r manifold.

Lemma 4.11. *For any $\epsilon \in (0, \epsilon^*]$, (4.27) has a C^r invariant manifold $M_\epsilon = W_\epsilon^+ \cap W_\epsilon^-$, which is the graph of a C^r function $\rho(\cdot, \cdot; \epsilon) : \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\}$. In addition, M_ϵ is the largest invariant subset of $\{\|p\| < \delta^*\} \times \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$.*

Proof. We consider an arbitrary, fixed $\epsilon \in (0, \epsilon^*]$ throughout this proof. Define the map $\mathcal{F}_\epsilon : \{\|p\| \leq \delta^*\} \times \{\|q\| \leq \delta^*\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|p\| \leq \delta^*\} \times \{\|q\| \leq \delta^*\}$ as follows:

$$\mathcal{F}_\epsilon(p, q, \zeta, \theta) := (\bar{\mathcal{P}}(q, \zeta, \theta; \epsilon), \bar{\mathcal{Q}}(p, \zeta, \theta; \epsilon)),$$

where $\bar{\mathcal{P}}(\cdot, \cdot, \cdot; \epsilon)$ and $\bar{\mathcal{Q}}(\cdot, \cdot, \cdot; \epsilon)$ are the continuous extensions of $\mathcal{P}(\cdot, \cdot, \cdot; \epsilon)$ and $\mathcal{Q}(\cdot, \cdot, \cdot; \epsilon)$ onto $\{\|q\| \leq \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$ and $\{\|p\| \leq \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$, respectively.

Consider the norm $\|(p, q)\|_1 := \|p\| + \|q\|$. By (4.35) and (4.37), we have that for any $(\zeta, \theta) \in \mathbb{R} \times \mathbb{R}^m$ and any (p_1, q_1) and $(p_2, q_2) \in \{\|p\| \leq \delta^*\} \times \{\|q\| \leq \delta^*\}$,

$$\begin{aligned} & \|\mathcal{F}_\epsilon(p_2, q_2, \zeta, \theta) - \mathcal{F}_\epsilon(p_1, q_1, \zeta, \theta)\|_1 \\ &= \|\bar{\mathcal{P}}(q_2, \zeta, \theta; \epsilon) - \bar{\mathcal{P}}(q_1, \zeta, \theta; \epsilon)\| + \|\bar{\mathcal{Q}}(p_2, \zeta, \theta; \epsilon) - \bar{\mathcal{Q}}(p_1, \zeta, \theta; \epsilon)\| \\ &\leq \frac{1}{2}\|q_2 - q_1\| + \frac{1}{2}\|p_2 - p_1\| \\ &= \frac{1}{2}\|(p_2 - p_1, q_2 - q_1)\|_1. \end{aligned}$$

Thus, for each $(\zeta, \theta) \in \mathbb{R} \times \mathbb{R}^m$, $\mathcal{F}_\epsilon(\cdot, \cdot, \zeta, \theta)$ is a contraction on $\{\|p\| \leq \delta^*\} \times \{\|q\| \leq \delta^*\}$ under the norm $\|\cdot\|_1$. Then we can define a function $\rho(\cdot, \cdot; \epsilon) : \mathbb{R} \times \mathbb{R}^m \rightarrow \{\|p\| \leq \delta^*\} \times \{\|q\| \leq \delta^*\}$ such that for each $(\zeta, \theta) \in \mathbb{R} \times \mathbb{R}^m$, $(p, q) = \rho(\zeta, \theta; \epsilon)$ is the unique solution to the equation

$$(4.39) \quad (p, q) = \mathcal{F}_\epsilon(p, q, \zeta, \theta).$$

Clearly, we have $\{(\rho(\zeta, \theta; \epsilon), \zeta, \theta) : (\zeta, \theta) \in \mathbb{R} \times \mathbb{R}^m\} = W_\epsilon^+ \cap W_\epsilon^-$, which verifies that M_ϵ is the graph of the function $\rho(\cdot, \cdot; \epsilon)$.

Next, we consider the restriction of \mathcal{F}_ϵ on $\{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \mathbb{R} \times \mathbb{R}^m$. By (4.36) and (4.38), we have that

$$(\bar{\mathcal{P}}(q, \zeta, \theta; \epsilon), \bar{\mathcal{Q}}(p, \zeta, \theta; \epsilon)) \in \{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\}$$

for all $(p, q, \zeta, \theta) \in \{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \mathbb{R} \times \mathbb{R}^m$. Thus, for each $(\zeta, \theta) \in \mathbb{R} \times \mathbb{R}^m$, $\mathcal{F}_\epsilon(\cdot, \cdot, \zeta, \theta)$ is also a contraction on $\{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\}$ under the norm $\|\cdot\|_1$, and the unique solution (p, q) to (4.39) is in fact contained in the set $\{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\}$. Therefore, $\rho(\cdot, \cdot; \epsilon)$ maps $\mathbb{R} \times \mathbb{R}^m$ into $\{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\}$.

Finally, we show that $\rho(\zeta, \theta; \epsilon)$ is C^r with respect to (ζ, θ) on $\mathbb{R} \times \mathbb{R}^m$ using the implicit function theorem. Define a C^r function $\mathcal{G}_\epsilon : \{\|p\| < \delta^*\} \times \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^{n_u} \times \mathbb{R}^{n_s}$ as follows:

$$\mathcal{G}_\epsilon(p, q, \zeta, \theta) := \begin{pmatrix} p - \mathcal{P}(q, \zeta, \theta; \epsilon) \\ q - \mathcal{Q}(p, \zeta, \theta; \epsilon) \end{pmatrix}.$$

Note that at any $(\rho(\zeta, \theta; \epsilon), \zeta, \theta) \in \{\|p\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \{\|q\| \leq 2\frac{C_3}{\sigma}\epsilon^\mu\} \times \mathbb{R} \times \mathbb{R}^m$,

$$D_{(p,q)}\mathcal{G}_\epsilon = \begin{pmatrix} I_{n_u \times n_u} & -D_q\mathcal{P} \\ -D_p\mathcal{Q} & I_{n_s \times n_s} \end{pmatrix}$$

is always invertible since $\|D_q \mathcal{P}\| \leq \frac{1}{2}$ by (4.35) and $\|D_p \mathcal{Q}\| \leq \frac{1}{2}$ by (4.37). Then, by the implicit function theorem, the function $\rho(\cdot, \cdot; \epsilon)$, which satisfies $\mathcal{G}_\epsilon(\rho(\zeta, \theta; \epsilon), \zeta, \theta) \equiv 0$ for all $(\zeta, \theta) \in \mathbb{R} \times \mathbb{R}^m$, is C^r on $\mathbb{R} \times \mathbb{R}^m$. \square

Finally, we return to the original system (4.20). By the periodicity of $\mathcal{P}(q, \zeta, \theta; \epsilon)$ and $\mathcal{Q}(p, \zeta, \theta; \epsilon)$ with respect to ζ and θ , we have that $\rho(\zeta, \theta; \epsilon)$ is $2\zeta_0$ -periodic in ζ and 2π -periodic in each component of θ . Recall the changes of variables (4.24) and $(a, b) = v(p, q)$. Then Lemma 4.11 implies that for any $\epsilon \in (0, \epsilon^*]$, (4.20) has an invariant set

$$T_\epsilon := \{(w, \theta) : w = \chi(\zeta) + \eta(\zeta) v(\rho(\zeta, \theta; \epsilon)), \zeta \in \mathbb{R}(\bmod 2\zeta_0), \theta \in \mathbb{T}^m\},$$

which is contained in an $\mathcal{O}(\epsilon^\mu)$ -neighborhood of T_0 since the function $\rho(\cdot, \cdot; \epsilon)$ is $\mathcal{O}(\epsilon^\mu)$. To show that T_ϵ can be parameterized by (4.23), which uses the ζ_0 -periodic basis $\hat{n}(\zeta)$, we take a section $\Sigma(\zeta_1)$ at an arbitrary $\zeta_1 \in [0, \zeta_0)$ as follows:

$$\Sigma(\zeta_1) := \{(w, \theta) : w = \chi(\zeta_1) + \eta(\zeta_1) v(p, q), \|p\| < \delta^*, \|q\| < \delta^*\}.$$

Clearly, we have

$$R(\zeta_1) := \{(w, \theta) : w = \chi(\zeta_1) + \eta(\zeta_1) v(\rho(\zeta_1, \theta; \epsilon)), \theta \in \mathbb{T}^m\} \subseteq T_\epsilon \cap \Sigma(\zeta_1).$$

Suppose there exist a $\zeta_2 \in [0, 2\zeta_0)$ different from ζ_1 and a $\theta_2 \in \mathbb{T}^m$ such that $(w_2, \theta_2) \in T_\epsilon \cap \Sigma(\zeta_1)$ for $w_2 = \chi(\zeta_2) + \eta(\zeta_2) v(\rho(\zeta_2, \theta_2; \epsilon))$. Since $(w_2, \theta_2) \in \Sigma(\zeta_1)$ and $\rho(\cdot, \cdot; \epsilon)$ is $\mathcal{O}(\epsilon^\mu)$, there exists (\tilde{p}, \tilde{q}) with both $\|\tilde{p}\|$ and $\|\tilde{q}\|$ being $\mathcal{O}(\epsilon^\mu)$ such that $w_2 = \chi(\zeta_1) + \eta(\zeta_1) v(\tilde{p}, \tilde{q})$. Furthermore, since $(w_2, \theta_2) \in T_\epsilon$ and T_ϵ is contained in an $\mathcal{O}(\epsilon^\mu)$ -neighborhood of T_0 , the solution trajectory of (4.20) that passes through (w_2, θ_2) is contained inside an $\mathcal{O}(\epsilon^\mu)$ -neighborhood of T_0 forever in both forward time and backward time due to the invariance of T_ϵ under the flow of (4.20). Then for the normal form (4.27), the solution trajectory that passes through $(\tilde{p}, \tilde{q}, \zeta_1, \theta_2)$ stays inside $\{\|p\| < \delta^*\} \times \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$ forever in both forward time and backward time. Recall that M_ϵ is the largest invariant subset of $\{\|p\| < \delta^*\} \times \{\|q\| < \delta^*\} \times \mathbb{R} \times \mathbb{R}^m$. Thus $(\tilde{p}, \tilde{q}, \zeta_1, \theta_2) \in M_\epsilon$, which implies that $(\tilde{p}, \tilde{q}) = \rho(\zeta_1, \theta_2; \epsilon)$. It follows that $(w_2, \theta_2) \in R(\zeta_1)$. Then we have $R(\zeta_1) = T_\epsilon \cap \Sigma(\zeta_1)$. Note that $T_\epsilon \cap \Sigma(\zeta_1)$ is C^r diffeomorphic to the intersection of T_ϵ and the section $\{(w, \theta) : w = \chi(\zeta) + \hat{n}(\zeta) \xi, \xi \in \mathbb{R}^{n-1}, \|\xi\| < \Delta_0\}$ for a certain $\Delta_0 > 0$ and any $\zeta \in [0, \zeta_0)$. Therefore, T_ϵ can be parameterized by (4.23), and it is the unique invariant torus for (4.20) inside an $\mathcal{O}(1)$ -neighborhood (i.e., independent of ϵ) of T_0 .

APPENDIX A. THE WAŻEWSKI PRINCIPLE

We follow the presentation of Conley [8]. Let \mathcal{X} be a topological space and $\varphi : \mathbb{R} \times \mathcal{X} \rightarrow \mathcal{X}$ be a flow. For a set $W \subset \mathcal{X}$, we define the following sets:

$$W^0 := \{x \in W : \exists t > 0 \text{ such that } \varphi(t, x) \notin W\},$$

$$W^- := \{x \in W : \varphi([0, t], x) \not\subseteq W \text{ for all } t > 0\},$$

where W^0 is the set of points that do not stay in W forever under the flow φ in forward time, and W^- is the set of points that immediately leave W in forward time. Clearly, $W^- \subseteq W^0 \subseteq W$.

Definition (Ważewski Set). The set W is called a Ważewski set if the following conditions are satisfied:

(W1) If $x \in W$ and $\varphi([0, t], x) \subset \text{cl}(W)$ then $\varphi([0, t], x) \subset W$.

(W2) W^- is closed relative to W^0 .

Theorem (Ważewski). *If W is a Ważewski set, then W^- is a strong deformation retract of W^0 and W^0 is open relative to W .*

The proof of the above theorem can be found in [8, 16]. Here we only recall that when W^- is a strong deformation retract of W^0 , there exists a continuous function $r : W^0 \times [0, 1] \rightarrow W^0$ such that: (1) for all $x \in W^0$, $r(x, 0) = x$ and $r(x, 1) \in W^-$; and (2) for all $x \in W^-$ and all $\sigma \in [0, 1]$, $r(x, \sigma) = x$. The function r is called a *strong deformation retraction*.

REFERENCES

- [1] V. I. ARNOLD, V. V. KOZLOV, AND A. I. NEISHTADT, *Mathematical Aspects of Classical and Celestial Mechanics*, vol. 3 [Dynamical Systems. III] of Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 2nd ed., 1997.
- [2] P. W. BATES AND C. K. R. T. JONES, *Invariant manifolds for semilinear partial differential equations*, in Dynamics Reported, vol. 2, Wiley, Chichester, 1989, pp. 1–38.
- [3] P. W. BATES, K. LU, AND C. ZENG, *Existence and persistence of invariant manifolds for semiflows in Banach space*, Mem. Amer. Math. Soc., 135 (1998), pp. viii+129.
- [4] N. N. BOGOLIUBOV AND Y. A. MITROPOLSKY, *Asymptotic Methods in the Theory of Non-Linear Oscillations*, International Monographs on Advanced Mathematics and Physics, Hindustan Publishing Corp., Delhi, Gordon and Breach Science Publishers, New York, 1961. Translated from the second revised Russian edition.
- [5] C. CHICONE AND W. LIU, *On the continuation of an invariant torus in a family with rapid oscillations*, SIAM J. Math. Anal., 31 (1999/00), pp. 386–415 (electronic).
- [6] S.-N. CHOW, W. LIU, AND Y. YI, *Center manifolds for invariant sets*, J. Differential Equations, 168 (2000), pp. 355–385.
- [7] S.-N. CHOW AND K. LU, *Invariant manifolds and foliations for quasiperiodic systems*, J. Differential Equations, 117 (1995), pp. 1–27.
- [8] C. CONLEY, *Isolated Invariant Sets and the Morse Index*, vol. 38 of CBMS Regional Conference Series in Mathematics, American Mathematical Society, Providence, R.I., 1978.
- [9] N. FENICHEL, *Persistence and smoothness of invariant manifolds for flows*, Indiana Univ. Math. J., 21 (1971/1972), pp. 193–226.
- [10] J. GUCKENHEIMER AND P. HOLMES, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, vol. 42 of Applied Mathematical Sciences, Springer-Verlag, New York, 1990. Revised and corrected reprint of the 1983 original.
- [11] G. HALLER, *Chaos Near Resonance*, vol. 138 of Applied Mathematical Sciences, Springer-Verlag, New York, 1999.
- [12] G. HALLER AND S. WIGGINS, *Geometry and chaos near resonant equilibria of 3-DOF Hamiltonian systems*, Phys. D, 90 (1996), pp. 319–365.
- [13] A. HARO AND R. DE LA LLAVE, *A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of hyperbolicity*, SIAM J. Appl. Dyn. Syst., 6 (2007), pp. 142–207 (electronic).
- [14] M. W. HIRSCH, C. C. PUGH, AND M. SHUB, *Invariant Manifolds*, vol. 583 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1977.
- [15] C. K. R. T. JONES, *Geometric singular perturbation theory*, in Dynamical Systems (Montecatini Terme, 1994), vol. 1609 of Lecture Notes in Mathematics, Springer, Berlin, 1995, pp. 44–118.
- [16] T. KACZYNSKI, K. MISCHAIKOW, AND M. MROZEK, *Computational Homology*, vol. 157 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.
- [17] N. KOPELL, *Invariant manifolds and the initialization problem for some atmospheric equations*, Phys. D, 14 (1985), pp. 203–215.
- [18] R. MCGEHEE, *The stable manifold theorem via an isolating block*, in Symposium on Ordinary Differential Equations (Univ. Minnesota, Minneapolis, Minn., 1972), vol. 312 of Lecture Notes in Mathematics, Springer, Berlin, 1973, pp. 135–144.
- [19] S. WIGGINS, *Normally Hyperbolic Invariant Manifolds in Dynamical Systems*, vol. 105 of Applied Mathematical Sciences, Springer-Verlag, New York, 1994.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853
E-mail address: `gy26@cornell.edu`